# Battery Safety Training in Virtual Reality

## Evaluating Virtual Reality as an Industrial Training Platform for the Future Battery Industry

Master's Thesis in Computer science and engineering

Marianne Garabetian & Nazmiyeh Sadiyeh

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2025

Master's Thesis 2025

# Battery Safety Training in Virtual Reality

Evaluating Virtual Reality as an Industrial Training Platform for the Future Battery Industry

Marianne Garabetian
Nazmiyeh Sadiyeh

UNIVERSITY OF GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2025

Battery Safety Training in Virtual Reality
Evaluating Virtual Reality as an Industrial Training Platform for the Future Battery Industry
Marianne Garabetian
Nazmiyeh Sadiyeh



Supervisor: Henrik Söderlund, Department of Industrial and Materials Science
Examiner: Gregory Gay, Department of Computer Science and Engineering

Master's Thesis 2025
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: An illustration of a woman wearing a VR headset and using hand-tracking.

The image was generated using Dall-E, a generative AI tool that can create images. The image was generated by prompting "a woman using VR in an educational task. Include elements like virtual warning signs or safety icons and highlight the VR technology. Include hand tracking mechanism and a white headset,"

Typeset in LaTeX
Gothenburg, Sweden 2025

Battery Safety Training in Virtual Reality
Evaluating Virtual Reality as an Industrial Training Platform for the Future Battery Industry
Marianne Garabetian
Nazmiyeh Sadiyeh
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

# Abstract

This thesis explores the potential of Virtual Reality (VR) as a platform for industrial safety training, using the emerging Swedish battery industry as a case study. With the increasing demand for a skilled workforce in hazardous environments, such as battery manufacturing, traditional training methods fall short due to limitations in realism, scalability, and safety. This research evaluates the effectiveness of VR training on knowledge retention, user engagement, and real-world task performance through empirical user studies conducted in collaboration with Battery Centre Gothenburg (BCG). A pre-study involving user testing with 78 participants, a literature search and stakeholder interviews laid the foundation for a detailed, validated requirement specification. The specification consists of 6 categories and 61 individual requirements. These requirements aimed to guide the design of effective VR training systems from a Human-Computer Interaction (HCI) perspective. A Proof of Concept (PoC) VR scenario was developed in Unity, in order to test and validate the results in the requirement specification. The PoC incorporated key HCI features, identified in the requirements specification, such as intuitive interaction, real-time feedback systems and cognitive load management. The requirement specification was evaluated through surveys and a measured real-world safety task, compared to a control group, with a total of 14 participants. Results show that the VR training improved recall of safety procedures, reduced task completion time and improved both user comfort and immersion. However, participants reported lower post-training confidence and risk understanding compared to prior training, indicating a need for deeper reflection and more complex, nuanced training scenarios. The discussion highlights the role of HCI in supporting engagement and usability, and the importance of balancing simplicity with instructional depth. The thesis concludes that VR can be an effective tool for safety training in high-risk industries when designed with user needs, technical feasibility, and learning goals in mind. A validated requirement specification and insights from real-world testing offer a foundation for future development of scalable and effective VR-based training systems.

Keywords: virtual reality (VR), safety training, battery safety, industrial training, knowledge retention, HCI, software engineering, requirements engineering, usability evaluation, VR development

# Acknowledgements

We would like to thank our supervisors Henrik Söderlund and Daniel Nåfors for their invaluable guidance, encouragement and feedback throughout the course of this thesis. Their expertise and advice were crucial to the development and completion of this research. We would also like to thank BCG for providing access to their facilities, systems, and expertise. A special thanks to the BCG instructors for their collaboration, feedback, and valuable insights, which greatly enriched the research. Our thanks also go to our examiner Gregory Gay and our opponents for reviewing the thesis and providing valuable feedback. Finally, we are deeply grateful to the participants who took part in the user testing and evaluations for their time and willingness to contribute to this study.

Marianne Garabetian, Gothenburg, June 2025
Nazmiyeh Sadiyeh, Gothenburg, June 2025

# Declaration of Generative AI

During the preparation of this work, ChatGPT was used in order to refine grammar and enhance clarity in phrasing. The tool did not generate substantial original content. The content was reviewed and edited as needed and the authors take full responsibility for the content of the published thesis.

Marianne Garabetian, Gothenburg, June 2025
Nazmiyeh Sadiyeh, Gothenburg, June 2025

# Contents

# List of Figures

# List of Tables

# 1 Introduction

In recent years, there has been a growing focus on battery production as a means to support the ongoing electrification in Sweden. This growth has created an urgent need for a skilled battery workforce, equipped with knowledge in areas such as battery cell manufacturing, battery assembly and recycling [1]. Today, this workforce is still evolving as the battery industry is taking its shape in Sweden and a talent shortage is evident [2].

In response to this need, the city of Gothenburg (Göteborgs Stad), in collaboration with Västra Götalandsregionen, conducted a pilot study that identified the need for specialized battery training [3]. As a result, an initiative was launched to establish a dedicated training center, which Gothenburg Technical College (GTC) was awarded the responsibility of designing, developing, and operating. This is now known as Battery Centre Gothenburg (BCG). BCG provides essential skills and safety training for the future battery workforce in Västra Götaland. One of the objectives of the center is to upskill and reskill employees to be able to safely and effectively transition to work in a battery factory. This includes training on chemical handling used in cell production, as well as electrical safety and working in high voltage areas. Education at BCG includes traditional classroom training, physical workstation training and various virtual training modules using extended reality technology [4].

This study aims to evaluate the effectiveness of Virtual Reality (VR) as a safety training tool in the context of Sweden's battery industry environment. Although the study focuses on the battery industry as a case study, the findings aim to contribute to a deeper understanding of how VR can support training for hazardous industrial environments. The research focuses on evaluating the effect of VR training on learning outcomes and usability by conducting an empirical study including participants in BCG's training program. Refinement and enhancements of the VR training methodologies will be derived from the findings. This thesis intends to contribute an assessment of the current VR training system, data on the effectiveness of learning and usability, and implement the proposed improvements based on the study outcomes.

There are several challenges emerging as VR is introduced as a training method in manufacturing, including its usability, effectiveness and inclusivity [2][5]. To ensure that the VR training system meets industry demands, it needs to be designed with human factors in mind [4]. Such challenges include ease of use and engagement, consideration for diverse skill levels and maintaining real-world applicability [2][5]. There is a need for a structured approach to design effective VR training solutions [6]. This study applies Human-Computer Interaction (HCI) and User-Centered Design (UCD) principles to ensure that the VR system is aligned with user and industry needs.

## 1.1 Problem Description

In modern industry, there is a growing demand for effective and scalable workforce training, driven by rapid technological advancements, increased safety regulations and the need to maintain high productivity levels in competitive markets [4][7]. This is especially important in physically demanding, high-risk, or safety-critical environments such as manufacturing, construction, biotechnology, and chemical processing, where errors can have serious consequences for both personnel and equipment [8]. Traditional training methods, such as classroom instruction and onsite practice, often fall short, as they are frequently constrained by safety risks, high costs, limited availability of expert instructors, and a lack of realistic practice environments [9][10]. Moreover, they frequently fall short in preparing workers to respond to real-life emergencies or hazardous scenarios [7].

VR offers a promising alternative by allowing trainees to practice procedures and handle emergency scenarios in immersive, interactive, and safe environments. Research shows that VR training can improve knowledge retention, engagement, hazard awareness, and task accuracy across several industries [2][11]. It supports repeated practice of hazardous tasks and can simulate stressful or rare situations without real-world risks [12][10]. However, using VR in training poses its own challenges. Designers must consider mental strain, how realistic and easy to use the simulation is and whether the experience actually helps in real-life situations [13][14]. The training system should also be adapted to the users' needs and match the complexity of the task [5].

This thesis examines these broader challenges with VR in industrial training through a case study in the Swedish battery manufacturing industry. Many companies are starting to explore and use VR training as a safer and more effective alternative to traditional training methods. This use case is well suited for this study, as it reflects common challenges faced in high-risk industries and contributes to solving broader problems related to the design and implementation of VR-based job training.

## 1.2 Aims of the Study

This study aims to investigate gaps in the application of VR training within the battery industry by evaluating its effectiveness in improving safety trainings and usability. By applying software engineering and HCI principles, the research seeks to refine VR training methodologies and propose enhancements that align with the specific needs of the industry. The findings aim to contribute to the broader understanding of the role of VR in industrial training by providing new insights to extend existing knowledge and offering a framework for designing scalable and inclusive training systems.

More specifically, the study focuses on evaluating the effectiveness of a VR training system that was developed and launched at BCG in January 2025. The study evaluates students' ability to learn and transfer knowledge using the simulated VR environment as well as how VR technology can facilitate learning. Based on these findings, the study aims to develop a validated requirement specification for VR safety training, grounded in HCI principles and user feedback. This specification serves as a framework to guide practitioners in designing VR environments that are effective, engaging and scalable for industrial training purposes. To validate this framework, a Proof of Concept (PoC) VR training scenario will be developed in Unity. This PoC will incorporate core HCI features such as intuitive interaction, real-time feedback and cognitive load balancing. Ultimately, the aim is to evaluate and improve the existing VR training system for future educational use at BCG, and to contribute insights that can inform safety training practices in other high-risk industrial sectors. Beyond the battery industry, the insights from this research aim to offer generalizable guidance and validated frameworks for designing effective VR-based training solutions across a range of safety-critical environments.

To achieve this, the study examines the following research questions:

**RQ1: What requirements need to be considered for VR to be effectively used as a training tool in industrial settings, with battery safety training as a use case?**
This question focuses on exploring the requirement fundamentals necessary to ensure the system's ability to address the needs of the industry and users, and to function effectively as a training tool in industrial safety training. The focus will be on general VR training needs while also including specific factors from the battery safety industry. Aspects such as functional requirements, non-functional requirements, system and data requirements and validation will be explored.

**RQ2: How can HCI principles impact engagement and effectiveness in VR-based training?**
This question aims to explore how a user-centered approach can improve the learning experience in VR training. It focuses on evaluating HCI strategies such as interface design, interaction feedback mechanisms and cognitive load influence, to optimize usability, engagement and the effective transfer of knowledge.

To address RQ1 and RQ2, this study will investigate how the current VR-based safety training system can be improved based on established requirements and HCI principles. Furthermore, it will explore how these insights can be directly implemented to evaluate and enhance the current VR training system. The study will determine usability gaps and implement improvements to optimize the learning experience.

## 1.3 Significance of the Study

This thesis contributes to both academic research and industry practice. It investigates the use of VR as a tool for safety training in the battery manufacturing industry. This is an area where VR applications are still limited and not well studied. The study follows a user-centered and iterative design approach. It is based on principles from HCI and software engineering. These principles help ensure that the training system is both effective and adapted to the needs of real users. This is important in high-risk industrial environments where traditional methods often lack realism, engagement, or flexibility.

For researchers, the thesis presents a structured and repeatable approach to designing VR training systems. It builds on established frameworks and uses empirical user feedback. While the focus is on battery production, the training challenges addressed in this study are common in many other high-risk fields. These include hazard awareness, complex tasks, and safety-critical decision-making. This means the approach could be useful in other industrial settings as well. For practitioners, the thesis offers clear guidance and useful methods. These can support the planning and development of VR-based safety training in real work environments.

## 1.4 Delimitations

This thesis is delimited by its focus on short-term learning outcomes. No long-term evaluation of knowledge retention or on-the-job performance is conducted. The study is also delimited to a specific user group consisting of students and blue-collar workers in the battery manufacturing industry. It does not include supervisors, managers or participants from other industrial sectors. The training system is assessed only in terms of usability and immediate learning effectiveness, without detailed investigation into emotional responses, long-term engagement or behavioral change. Finally, the study uses the existing VR hardware setup provided by BCG, without any technical modifications or enhancements.

## 1.5 Thesis Outline

This thesis is organized as follows. Chapter 2 provides a comprehensive background of VR, including its key concepts, history, development, ethical considerations and use as a training tool. The background chapter also covers learning theories, HCI principles and an overview of battery production, which is important for the use case. Chapter 3 provides findings from previous case studies related to VR training and its effectiveness. Chapter 4 describes the methodology used to conduct the research, followed by chapter 5 presenting the results. Finally, chapter 6 provides a discussion and chapter 7 presents the conclusion.

# 2

# Background

This chapter presents a comprehensive overview of VR, its historical development, its evolution as a training tool and ethical considerations. It further explores the software development methodologies relevant to the study and VR's growing role in industrial workforce training. Moreover, HCI principles associated with immersive environments are explored. Finally, this chapter provides an overview of battery production, which is relevant to the use case.

## 2.1 Key Concepts of VR

VR is a technology that creates simulated environments. VR allows users to experience and engage with digital spaces in a way that feels immersive and realistic. This interaction creates a dynamic and responsive experience. These technologies provide sensory input that simulates real-world experiences, such as sight, sound, and touch [15]. However, VR is part of a broader spectrum of immersive technologies, known as the Virtuality Continuum. This continuum ranges from the real world to fully virtual environments, with Augmented Reality (AR) and Mixed Reality (MR) in between. AR adds digital objects to the real world, as in mobile AR apps and smart glasses, while MR blends real and virtual objects to interact in real time, which is used in Microsoft HoloLens. Fully immersive VR, on the other hand, completely surrounds the user in a computer-generated world using headsets and motion tracking, blocking out the real environment [16].

To create an immersive and convincing VR experience, three key principles come into play: Immersion, Interaction, and Presence. Immersion refers to the user's sense of being fully absorbed by the virtual environment. It is achieved through visuals and sensory feedback. Interaction is the ability to manipulate virtual objects or navigate the environment. It allows users to move through and control objects in the virtual space using controllers, hand gestures or tracking systems. Presence is the psychological sensation of "being there" in the virtual world. This principle is crucial for creating convincing VR experiences. It helps users emotionally connect with the environment, making the experience more engaging and impactful. The stronger the sense of presence, the more immersive and realistic the VR experience becomes [15].

### 2.1.1 VR Foundation and Future Outlook

The creation of VR started in the 19th century. Immersive art forms like panoramic murals and stereoscopic viewers aim to create a fully engaging visual experience. Early examples of immersive experiences included stereoscopic viewers, which created the illusion of depth. It also included panoramic murals, which were large-scale paintings or images that wrapped around a space to create the illusion of being inside a scene. However, the modern VR technology that we know today actually started in the 1960s. In 1962, Morton Heilig introduced the Sensorama, which was the first multi-sensory immersive system. This system simulated a motorcycle ride through New York, see Figure 2.1. It had 3D visuals, sound, motion and even scent to create a fully immersive experience [17][18].

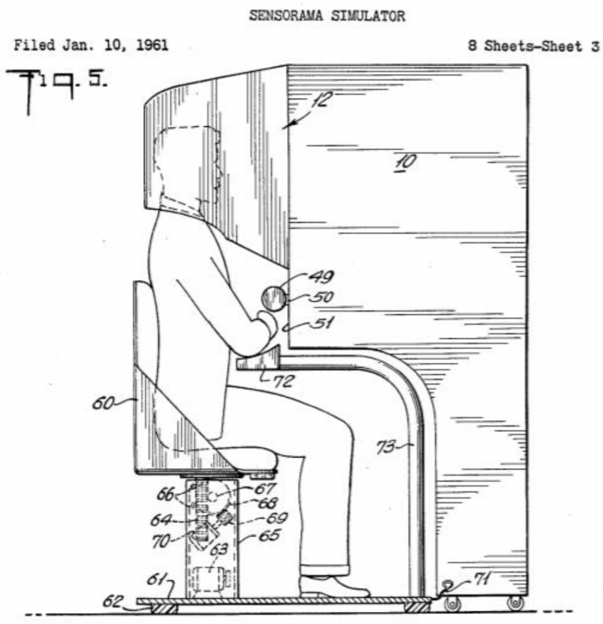



**Figure 2.1:** Patent illustration of the Sensorama Simulator developed by Morton Heilig in 1962. Taken from [19]

Ivan Sutherland is a pioneer in computer graphics. He introduced the "ultimate display" in 1965. This system combined interactive visuals, spatial sound, haptic feedback, and multiple senses like sound, smell, and taste. His concept aimed to create a simulated environment, a computer-generated space that mimics real or imagined settings. His idea led to the development of the first head-mounted display (HMD) in 1968, in collaboration with David Sproull. An HMD is a wearable screen placed in front of the eyes to display virtual images. It allowed users to see a digital environment instead of the real world. The device featured stereoscopic visuals, a technique that creates the illusion of depth. It did this by showing slightly different images to each eye. This mimicked how human vision naturally perceives depth. The HMD also tracked head movements, allowing real-time adjustments to the virtual world as users moved their heads. This innovation was a breakthrough in VR technology [17][18].

From the 1960s to the 1990s, VR technology advanced significantly due to improvements in computer graphics, processing power, and simulation systems. In the 1980s, VR was used in the military, especially for pilot training. The U.S. Air Force created the Visually Coupled Airborne Systems Simulator (VCASS), which gave pilots real-time data to help them better understand their surroundings during simulated combat and provided interactive decision-making tools. This technology was important for military training, where realism and fast decision-making were crucial [17].

Around the same time, NASA saw the potential of VR for astronaut training and remote controlled robotics. One of their projects was the Virtual Wind Tunnel. It was a system that let scientists study how air flows around spacecraft and airplanes without needing a physical wind tunnel. This made testing safer, cheaper, and much more flexible. NASA also developed remote telepresence systems, allowing astronauts to control robotic devices in hazardous environments. Imagine operating a robot on Mars while sitting safely on Earth, which VR made possible [17][18].

In addition to simulations, NASA explored human-computer interaction (HCI) advancements through another project. NASA developed *The Virtual Interactive Environment Workstation (VIEW)* in 1988, see Figure 2.2. The system had integrated hand tracking, speech recognition, and spatial 3D audio. This system laid the foundation for modern gesture-based and voice-controlled VR interfaces that are widely used today. The hand tracking led to gesture based interactions, reduced reliance on controllers and allowed users to pick up and manipulate virtual objects. The speech recognition allowed hands-free navigation in VR simulations, while the 3D audio improved the situational awareness, which helped users perceive directional audio cues from the environment [17][20].


**Figure 2.2:** NASA's Virtual Interactive Environment Workstation (VIEW) was developed in 1988. Taken from [21].

In the 1990s, VR became available to the public with systems like Virtuality. This

VR arcade game allowed users to explore 3D virtual worlds. Although these early systems had technical limitations, they got the public interest. VR technology then expanded into fields like scientific research, medical training, manufacturing training and entertainment [18][2]. As VR improved, it began to be used for more than just entertainment. It was especially helpful in training and education, where its immersive features enhanced learning [15].

#### Future of VR

The future of VR is shaped by new technologies that improve visual quality, comfort, and user interaction [22]. Innovations like holographic optical elements (HOEs), metasurfaces, and micro-LEDs will make VR headsets lighter, clearer, and more energy efficient. At the same time, they still offer high resolution and a wide field of view. Foveated rendering and varifocal displays are new techniques that improve user comfort. Foveated rendering saves processing power by making only the area the user is looking at sharp, while the rest stays slightly blurred. Varifocal displays change focus automatically to make visuals look more natural and reduce eye strain. Future VR systems will include gesture tracking and realistic touch feedback. It will also include brain-computer interfaces (BCI). However, VR still faces high development costs, interaction limitations, and ongoing concerns about privacy and digital addiction that need to be addressed. Ongoing research in display technology and user experience improvements is expected to make VR more comfortable, accessible, and widely used across different industries [22].

### 2.1.2 Components of Modern VR

Modern VR systems rely on a combination of hardware and software components to fulfill the three VR key principles: Immersion, Interaction, and Presence. These components enhance realism, usability, and functionality.

The advancements in hardware primarily determine the effectiveness of VR systems, which define how immersive and interactive the experience can be. The main hardware components included the following:

- Head-mounted displays (HMDs)
- Motion tracking
- Haptic feedback
- 3D Spatial audio
- Input devices

The HMD is one of the most important parts of VR. HMDs are the main way users see and interact with the virtual world. HMDs show 3D images that create a sense of depth and a wide field of view (FOV) to make the experience feel real. Organic Light-Emitting Diode (OLED) or Liquid Crystal Display (LCD) screens can be used in the HMD. The screens need to refresh at rates above 90Hz to minimize motion sickness. Popular VR headsets, such as the Meta Quest, HTC Vive, Valve Index and Varjo XR-3, use built-in or external sensors to track movement accurately, which helps users move naturally in the virtual space [18][14].

Hand-tracking technology has become an important way to interact in VR. NASA's VIEW system used gloves with sensors to let users control virtual objects naturally. Today, VR systems use advanced finger tracking and gesture recognition to make interactions even more realistic. Nowadays, technologies like Ultraleap and Meta's hand-tracking software let users control VR with just their hands, removing the need for controllers [20]. A big improvement in VR is haptic feedback, which is a technology that simulates the sense of touch in a virtual environment. This technology enables users to feel vibrations or pressure, making the experience more realistic. In recent years, VR-systems have small motors built-in in the controllers to create vibrations when touching or moving objects. More advanced haptic gloves and bodysuits, like the TESLASUIT, have been developed, which take haptic feedback into force feedback. Users are able to feel textures, resistance and the weight of objects, making the interaction even more realistic than before. Haptic feedback is relevant in medical training, industrial simulations and military exercises [14].

### 2.1.3 Ethical Considerations

As VR technologies become more widely used for training, several ethical considerations must be addressed to ensure responsible development and deployment. A significant ethical concern in VR training revolves around the potential for manipulation and the subsequent gradual loss of user independence. The immersive nature of VR can indirectly shape users' perceptions, emotions, and cognitive processes, often without their full awareness. In educational settings, highly realistic VR scenarios can subtly shape students' understanding of concepts, sometimes leading to bias or misinformation. Similarly, in corporate training, companies might use VR to condition employees to accept certain behaviors, policies, or ideologies without critical evaluation [23].

VR training environments collect large amounts of personal data. This includes eye movement, attention patterns, body language, and interaction behavior. In education, this data helps assess student engagement and understanding. However, if misused by institutions or third parties, it poses privacy risks [24][23]. In professional training, companies may track employees' cognitive responses, performance, and adaptability. This data could influence hiring, promotions, or terminations. Without strong data protection, VR training may become a tool for surveillance rather than skill development [23].

Accessibility is a major ethical concern in VR training. While VR enhances learning, not all users benefit equally. Research shows that individuals with low spatial ability may struggle with highly realistic environments, making training less effective for them. Those vulnerable to cybersickness or with neurological impairments may find VR training difficult or even harmful. If organizations rely only on VR training without alternatives, they risk excluding employees or students who cannot fully engage with the technology. Ethical training programs must be inclusive by offering multiple learning formats [24].

Prolonged use of VR training can lead to both physical and psychological side effects, including reduced autonomy and stress caused by the highly realistic environment. Users may experience motion sickness, eye strain, and fatigue, which can negatively affect learning and performance [24][23]. In high-risk fields, such as military or emergency response training, VR simulations can also induce psychological distress when users are exposed to highly realistic and traumatic situations. To reduce these risks, ethical VR design should include adjustable realism settings, content warnings, and regular breaks to protect the mental well-being of trainees [23].

## 2.2 VR Development

This section provides a brief and general description of VR development in Unity and relevant software development methodologies to use as a framework for the development of the PoC.

### 2.2.1 Introduction to Unity

Unity [25] is one of the most popular game engines for VR development. It is known for its utility in creating 2D and 3D games and applications in VR, AR and MR. Unity Technologies introduced Unity in June 2005 at Apple's Worldwide Developers Conference (WWDC) as a Mac-exclusive tool, but over the years, it has become a cross-platform development engine supporting over 20 platforms. Today, Unity is not only used in the gaming industry, but also in industries like automotive design, engineering, film and architecture [26]. In addition, Unity supports several VR platforms, such as Oculus, PlayStation VR, HTC Vive and Windows Mixed Reality, which makes it easy for developers to create VR applications that will work across multiple devices without having to rewrite the code for each platform. Unity also provides built-in tools to handle motion tracking, spatial audio and real-time physics, which are crucial for an immersive VR experience [26]. Unity Asset Store [27] offers a detailed library of pre-made VR assets. These include 3D models, shaders (which define the appearance of 3D models), and scripts, speeding up the development process [28]. Performance is an important factor in VR development, considering low frame rates can cause motion sickness. In Unity, there are optimization tools and dynamic resolution adjustments that can be used to ensure smooth performance, even on hardware-limited VR systems [26].

Unity makes it easier to build VR experiences. It uses the XR Interaction Toolkit, which includes tools like XRController, XRGrabInteractable, and Teleportation-Area. These let users grab objects, press buttons, and teleport in virtual environments. Unity also works well with other systems like SteamVR and Oculus Integration and integrates well with third-party SDKs such as SteamVR and Oculus Integration, allowing developers to tailor their applications to specific headsets and enhance compatibility across devices [26][25]. To speed up development, Unity's Asset Store provides a wide selection of ready-to-use assets such as VRTK, environment packs, UI systems, shaders, and scripting tools. These save time by reducing the need to build everything from scratch [28][27].

### 2.2.2 Software Development Methodologies

There are several frameworks that can be utilized in order to ensure successful software development. Different methodologies are used depending on the project scope and requirements. The most commonly used software development methodologies are presented below.

To address the problems with traditional methodologies, such as inflexibility, the Agile software development methodology emerged. Agile is an iterative process that focuses on adaptability and continuous feedback. The development process consists of sprints, which are short cycles where small functional parts of the software is delivered iteratively [29]. The key advantages of Agile are improved responsiveness when requirements change, increased team collaboration and fast delivery of functional software. By using Agile methodologies, customer feedback can be continuously incorporated, improving the final product [30]. However, Agile may be challenging for larger teams or companies with strict regulations and requirements, since it requires self-organized teams, frequent communication and continuous testing [31].

The Rapid prototyping approach focuses on creating a working prototype of a system quickly, before fully developing the system. This makes it possible to gather early user feedback and improve the design iteratively. The Rapid prototyping process includes the following phases: requirements, gathering, prototype development, user feedback collection, refinement and final product development [32]. The advantages of this method include risk reduction and user-centered design. Issues and mistakes are identified and addressed early in the process, before the full-scale development, which saves time and resources [32]. The continuous user feedback ensures that the final product meets user needs, improving the overall usability and effectiveness [30]. The drawback of Rapid prototyping is the risk of scope creep, which essentially means that continuous refinements and addition of new features can expand the project time and scope, leading to increased costs [32].

## 2.3 Workforce Training in Manufacturing Industries

In industrial sectors where safety and precision are critical, such as manufacturing, construction and chemical processing, training plays a central role in workforce preparation. Workers in these environments are often required to perform complex, high-risk tasks that demand theoretical understanding, procedural accuracy and situational awareness [12]. The effectiveness of training directly influences both operational performance and employee safety [33].

Traditional training methods, including classroom lectures, instructional manuals, and supervised on-the-job training (OJT), have long been the standard in industrial learning environments [33]. Classroom formats tend to rely heavily on passive knowledge transfer, limiting engagement and reducing retention, while OJT, though more experiential, can be logistically complex and limited by safety regulations, equipment availability, and trainer capacity [14].

To further address the challenges with traditional training methods, industries have progressively investigated digital solutions such as VR. VR-based training provides an interactive and immersive environment where trainees can practice real-world tasks in a realistic but risk-free setting. This makes hands-on learning possible without constraints such as physical equipment and safety risks [34].

An effective VR training system generally consists of four main components:

- **Input devices:** this could be controllers and/or hand-tracking sensors to simulate real-world interactions.
- **Output devices:** this is the VR headset, which provides an immersive visual experience.
- **Graphical systems:** this generates real-world simulations of for instance industrial processes.
- **Databases:** these store various scenarios, conditions, and training modules. In some cases, this includes learning management systems to track user progress and performance data [34].

One of the main advantages of VR training is that it allows for active participation, which leads to improved procedural memory and skill retrieval. Industries including construction, manufacturing and chemical processing are employing VR training to utilize cost-effective, scalable and adaptive learning experiences [14]. However, there are still challenges with implementing VR training in industries. There is a high initial investment cost, technical challenges and a need for specific knowledge in VR development [12]. In addition, factors such as motion sickness and cognitive overload can limit the learning effectiveness if the VR environment is not designed well [10]. This is further elaborated on in Section 2.5.

## 2.4 Learning Theories

Understanding how individuals develop skills and retain information is crucial for designing effective training systems. This section covers an overview of learning frameworks that can help evaluating the effectiveness on different training methods. In this study, the theories also serve as guiding frameworks for designing a VR training system that encourage effective learning.

### Experiential Learning Theory (ELT)

Kolb [35] defines learning as *"the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience."*. This emerged into the Experiential Learning Theory (ELT). ELT suggests a four-stage learning cycle:

- Concrete Experience (CE): this is when the learner directly interacts with a task or problem.
- Reflective Observation (RO): this is when the learner reflects on their experience.
- Abstract Conceptualization (AC): this is when the learner forms theories or conclusions based on the experience.
- Active Experimentation (AE): this is when the learner applies the knowledge to new situations.

### Nonaka & Takeuchi's SECI model

Nonaka & Takeuchi [36] developed the SECI (Socialization, Externalization, Combination, Internalization) Model, where they explain how organizations create, share and apply knowledge effectively. The model emphasizes the conversion between tacit knowledge (e.g. skills learned through experience) and explicit knowledge (e.g. documented, structured information) to improve learning and innovation. The model suggests four key processes for knowledge transfer. The first process is socialization (tacit-to-tacit), which refers to knowledge being transferred through shared experiences, such as trainer to trainee and hands-on practice. The second one is externalization (tacit-to-explicit), which refers to tacit knowledge being translated into explicit forms, such as manuals or documentation. The third process is combination (explicit-to-explicit), referring to when different parts of explicit knowledge are integrated to form new structured frameworks (e.g. safety protocols for industries). The fourth and last process is internalization (explicit-to-tacit), which refers to when explicit knowledge is consumed by individuals, converting it into personal expertise through repeated practice.

## 2.5 HCI Principles

HCI has a critical role in VR training, ensuring that the system is not only immersive, but also usable, engaging and more effective for learning. This section presents an overview of HCI principles, challenges, and user experience considerations based on findings from research.

### 2.5.1 Overview of HCI and UCD

HCI is a field focused on designing and improving how people interact with technology. Its goal is to make systems easier to understand, more efficient and accessible to all users. HCI is important for increasing productivity and reducing errors. It ensures technology is enjoyable and practical. It combines knowledge from psychology, computer science and design to create user-friendly interfaces. Applications range from websites and apps to virtual reality and industrial systems. By prioritizing usability, accessibility and adaptability, HCI ensures that technology meets the evolving needs of diverse users. UCD, which is part of the HCI field, is an iterative design process where the user is involved in the design process. It uses techniques like interviews, surveys and testing to determine their requirements and collect feedback [37]. This can benefit VR training by improving usability and learning, by using simpler interfaces to reduce cognitive overload and providing real-time feedback/reaction to the user's task [38]. Furthermore, incorporating features to adapt to users with different abilities and skill levels ensures that all workers can benefit from VR training [39].

### 2.5.2 Key HCI Principles and Considerations

Listed below are several HCI principles that need to be considered when designing VR training environments. These principles are important for ensuring smooth user experiences and learning outcomes.

**Presence and Immersion**
Presence is described by the user's sense of being physically present in the virtual environment. High presence leads to stronger learning engagement, but it is important to find a balance to avoid sensory overload for the user. To achieve high presence and immersion, factors such as field of view, haptic and audio feedback and natural user interaction with the environment are significant [40].

**Affordance**
Affordance is the user's ability to understand and interact with objects in virtual environments. In training contexts, affordance relevant to the task need to be built so that users can use them intuitively with little instructions or thinking. If objects in the virtual environment claim their purpose and role well, trainees are able to learn and use their knowledge in practice more effectively [38].

**Feedback Systems**
It is crucial in VR training to provide real-time feedback, as delays beyond 300ms have been shown to negatively impact user performance, sense of agency, and immersion [41]. There are several feedback types that can be included to create an effective feedback system. Visual feedback, could be for instance that interacted objects are highlighted. Audio feedback may include sound cues for warnings or approval and haptic feedback could be for instance vibrations when pressing a button [38].

**Cognitive load**
Another important aspect regarding VR training is the handling of cognitive load. The level of complexity of a virtual environment must be neither too simple nor too detailed. This ensures that users are not overwhelmed by excessive stimuli or unnecessary interactions. High visual density occurs when too many elements compete for the user's attention at once, which can reduce the learning efficiency and increase cognitive fatigue. Studies indicate that users perform tasks and recall information better when being in a structured environment, that balance detail with simplicity [39].

**Accessibility and Inclusivity**
It is important that VR training is adaptable for different types of users, considering their different abilities. One aspect is considering alternative interactions methods, for instance controller-based, hand-tracking or voice tracking interactions. Other aspects may be customization settings for the interface and providing VR training systems that can be used both seated and standing [39].

**Navigation**
Lastly, navigation is necessary for the user to maintain their orientation in the VR environment. Poor navigation can result in user disorientation and motion sickness. Some practices to include for VR navigation are teleportation-based movements (to reduce motion sickness) and visual markers to navigate in the environment (to improve spatial awareness) [39].

**Usability**
Usability is an aspect that needs to be evaluated to ensure that HCI principles have been implemented effectively. According to the ISO 9241-11 standard, usability is defined as *"the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"* [42]. This means that usability is not only about how well a system performs technically, but also how well it supports the user in achieving their goals comfortably and reliably. Listed below are some key usability measures that can be considered in VR training:

- **Task Completion Time:** measures how efficiently users complete the tasks. Lower completion times could indicate better usability [43].
- **Error Rate:** provides the number of mistakes made by a user during the training. Fewer errors could indicate better interaction design [43].
- **Engagement Score:** evaluates user focus and learning immersion when training with VR. Higher score could indicate better engagement [40].
- **Cognitive Load Score:** measures the mental effort needed to complete the training. Lower score could indicate a more efficient learning experience [39].
- **User Satisfaction Rating:** measured by surveys or interviews to evaluate how comfortable and intuitive users find the VR training experience [38].

## 2.6 Requirements Specification Framework

To guide the development of effective VR training systems, it is essential to adopt a structured requirements specification framework. In software and system development, a requirements specification framework serves as a structured method for capturing, organizing, and validating the needs and constraints of stakeholders. It bridges the gap between the problem and the technical implementation by translating user goals into actionable and testable software requirements. Clearly specified requirements play an important role in guiding design decisions and reducing misconceptions during the development process. They also help to ensure that the final product align with its intended purpose and user expectations [44].

Commonly used approach is based on the IEEE 830–1998 standard [45] for software requirements, which promotes a clear structure for software requirements specification (SRS) that distinguish the needs of different categories. Another widely referred framework is proposed by Lauesen [44], which includes landscape-based and target-oriented styles to better catch real-world references, especially in the system with a strong HCI component. The outline of Lauesen is divided into four main categories of requirements: functional, data, quality, and managerial requirements. The functional requirements specify what the system should do. It includes its characteristics, behavior and interaction. Data requirements define the structure and handling of information, including internal data representation. Quality requirements, traditionally referred to as non-functional requirements, specify how well the system should perform. These include aspects such as purposeful, performance, reliability and stability. Managerial requirements relate to project-level constraints such as delivery timelines, legal obligations, pricing models, and contractual responsibilities [44].

## 2.7 Battery Production

Over the decades, battery technology has evolved dramatically to meet the growing demands of efficiency, durability, and sustainability. There is a wide range of different batteries based on which chemical it is created from. Most of the commonly used battery types are lead-acid, nickel-based, lithium-ion, sodium-ion and emerging post-lithium battery technologies. However, lithium-ion batteries (LIBs) are the dominant choice for energy storage due to their high energy density, long cycle life, high efficiency and scalability. The production process of LIBs must be controlled to prevent contamination, ensure material uniformity and maintain cell stability. It is important to guarantee reliable and high-performance batteries [46].

The manufacturing of LIBs consists of an advanced process that involves multiple fundamental steps: raw material preparation, electrode manufacturing, cell assembly, formation and aging, and module & pack integration, see Figure 2.3. Each stage requires strict control over material purity, environmental conditions, and precision engineering to prevent contamination and ensure battery lifespan [47].

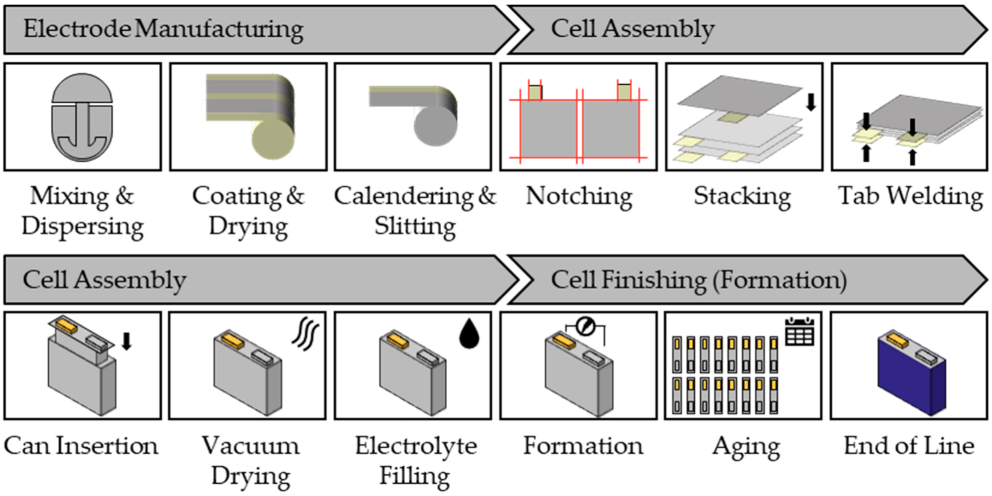



**Figure 2.3:** Production steps in lithium-ion battery cell manufacturing. Taken from [47].

The battery production process is automated. This is especially important because the process involves handling hazardous materials and requires strict quality control. However, workers must, from time to time, intervene when faults occur. These issues can include short circuits, leaking electrolytes, or overheating, which can be very risky. To handle these situations safely, workers need proper training. Many battery safety failures come from manufacturing defects that require human intervention to diagnose and fix [48].

To better understand where safety-critical tasks occur and where VR training can be applied, the production process can be divided into two main phases: Raw to Cell and Cell to Pack.

### Raw to Cell

This phase begins with raw material preparation. Here, key components such as the cathode, anode, electrolyte and separator are produced. These materials form the basic structure of a battery cell. Next comes electrode manufacturing, which is done in several steps. The raw materials are first mixed into a slurry. The slurry is then coated onto metal foils. After coating, the sheets are dried and pressed to a uniform thickness. This step is known as calendering. Finally, the foils are cut to the desired size. Once the electrodes are ready, they move into cell assembly. This stage involves stacking or winding the electrodes with a separator, filling the cell with electrolyte, and sealing the package. Precision is crucial here. Misaligned components or incorrect amounts of electrolyte can lead to performance issues or safety risks. After sealing, the cells go through formation. This step includes the first charging and discharging cycles, which create a stable internal layer called the solid electrolyte interphase (SEI). The final step is aging, where cells rest under controlled conditions to detect early defects before further use [47].

### Cell to Pack

Once individual cells are tested and approved, they are grouped into modules and eventually into full battery packs. First, the cells are sorted by key characteristics like capacity and resistance to ensure consistent performance. Next, the cells are connected electrically using methods such as laser welding or busbars. After that, additional systems are added to complete the pack. These include a Battery Management System (BMS), cooling components, and protective casings. Together, these systems help regulate temperature and ensure safe operation. Before use, battery packs undergo rigorous testing. Electrical tests check voltage and capacity, thermal tests evaluate heat resistance and safety tests identify risks such as short circuits or mechanical damage [47].

### Safety Considerations in Battery Manufacturing

When manufacturing a battery, safety is a top priority. Battery manufacturing involves the use of hazardous chemicals, high temperatures, and complex machines. Following safety measures helps reduce accidents, keep workers safe and ensure the battery quality. Table 2.1 provides an overview of the key safety risks in battery manufacturing along with the corresponding preventive measures [47].

| Safety Issue | What is the Risk? | How to Prevent It? |
|---|---|---|
| **Chemical Hazards** | Harmful chemicals like lithium salts and solvents can be toxic if inhaled or touched. | Use gloves, masks, protective clothing, good ventilation, and safe chemical handling procedures. |
| **Heat and Fire Risks** | Battery components can reach high temperatures, causing fires or explosions, especially if damaged or overheated. | Use cooling systems, monitor temperatures, install fire safety systems, and avoid overheating. |
| **Machine Hazards** | Workers may get hurt from sharp edges, heavy equipment, or moving parts in the production line. | Install safety barriers, use automated machines for dangerous tasks, and provide worker training. |
| **Electrical Hazards** | Batteries and equipment use high voltage, which can cause electric shocks or short circuits. | Use insulated tools, follow safe handling rules, wear protective gear, and have emergency shutdown systems. |
| **Contamination Issues** | Dust, dirt, or tiny impurities can lower battery quality and even cause failures or safety problems. | Maintain cleanrooms, check material purity, control contamination, and do strict quality testing. |

**Table 2.1:** Safety considerations in battery manufacturing.

Among the critical incidents in battery manufacturing, one of the most dangerous occurs when a battery cell is dropped, potentially releasing colorless and toxic gases like hydrogen fluoride (HF). Hydrogen fluoride is highly corrosive and flammable, posing life-threatening risks even at low concentrations. Table 2.2 shows the hazard classifications for hydrogen fluoride, based on international Globally Harmonized System (GHS) standards [49].

While automated systems reduce many risks, critical failures still require human intervention [48] [47]. These moments often occur under pressure, where improper handling of chemicals or PPE can escalate hazards, which underscores the need for targeted and realistic safety training [1].

| **Hazard Class** | **GHS Code** | **Hazard Statement** |
| --- | --- | --- |
| Gases under pressure, liquefied gas | H280 | Contains gas under pressure; may explode if heated |
| Acute toxicity, oral, category 2 | H300 | Fatal if swallowed |
| Acute toxicity, dermal, category 1 | H310 | Fatal in contact with skin |
| Skin corrosion/irritation, category 1A | H314 | Causes severe skin burns and eye damage |
| Serious eye damage/eye irritation, category 1 | H318 | Causes serious eye damage |
| Acute toxicity, inhalation, category 2 | H330 | Fatal if inhaled |
| Specific target organ toxicity, single exposure, respiratory tract irritation, category 3 | H335 | May cause respiratory irritation |
| Simple asphyxiant | H380** | May displace oxygen and cause rapid suffocation |

*Note:* **H380 is not an official GHS code.

**Table 2.2:** Reconstructed summary of hazard classifications and corresponding GHS hazard statements for hydrogen fluoride. Adapted from [49].

# 3

# Related Work

This chapter reviews previous research related to VR training in industrial contexts. It first reviews traditional training methods, highlighting their common practices and limitations in high-risk work settings. The chapter then explores case studies and empirical studies of VR-based training across various sectors. Finally, the chapter identifies gaps in current research and explains how this thesis contributes to advancing VR training methods for hazardous industries, particularly within the context of battery manufacturing safety training.

## 3.1 Traditional Training Today

Employee training is a key factor in enhancing job performance and maintaining organizational competitiveness in the modern workforce. According to Martin et al. [33], training involves the systematic development and delivery of knowledge, abilities, skills and attitudes (KASAs) intended to improve employee performance. Traditionally, organizations have relied on methods such as classroom lectures, printed manuals, workshops, and supervised on-the-job training experiences. These approaches have historically been preferred for their practicality, cost-effectiveness and ability to adapt to different training contexts and environments.

Lectures and classroom-based sessions are the leading traditional formats, focusing on theoretical knowledge transfer through instructor-led presentations. While lectures are efficient for communicating information to large audiences, they tend to promote a passive learning environment with limited learner engagement. Printed manuals and self-study documents further support this passive approach, requiring self-motivation and advanced literacy skills from employees. Such methods often fail to adapt to diverse cognitive styles and prior knowledge levels, impacting overall training effectiveness [33].

Burke et al. [50] conducted a meta-analysis of 95 studies to evaluate the effectiveness of various worker safety and health training methods. They categorized training into three levels of engagement: least (e.g., lectures), moderate (e.g., feedback-based), and most engaging (e.g., hands-on training). Their findings showed that while all methods produced positive effects, training that involved greater trainee participation significantly improved both knowledge acquisition and behavioral safety outcomes. For instance, the most engaging methods were nearly three times more effective than passive methods in increasing safety knowledge.

Theoretical perspectives from educational psychology also highlight the limitations present in traditional approaches. Kolb's ELT, which is previously introduced in Section 2.4, states that effective learning cycles require engagement across concrete experience, reflective observation, abstract conceptualization, and active experimentation. Traditional methods such as lectures and printed manuals primarily support abstract conceptualization and reflective observation, but often neglect active experimentation and concrete experience, which are critical for deep skill development and long-term retention [35]. Makransky and Petersen [13] argue that immersive learning environments help trainees learn by increasing their sense of presence and control, which boosts motivation, interest, and other mental processes that aren't usually activated in more passive learning settings.

New teaching approaches are starting to combine traditional methods, like classroom instruction, with newer technologies such as online tools and interactive learning strategies. Orey [51] describes how methods like experiential learning, cognitive tools, and computer-based instruction can make learning more engaging and flexible. These approaches aim to improve how learners interact with the material and help them apply what they learn in real situations.

## 3.2 VR-Based Training in Industrial Settings

Previous research on VR-based training highlights its effectiveness in enhancing learning retention, engagement, and skill acquisition. Studies from various industries have validated VR training using different methodologies and performance metrics.

In the construction industry, Osti et al. [9] developed a VR-based training system focused on teaching novice workers how to assemble wooden light-frame (WLF) structures. The study compared two groups: one trained using immersive VR tutorials, and the other using traditional 2D instructional videos. The results indicated clear advantages for the VR group, who completed the training 40% faster, committed 25% fewer assembly errors, and showed improved spatial understanding, which is an essential skill for structural assembly tasks. Participants also reported higher engagement and satisfaction with the training experience, suggesting that VR's immersive qualities foster better learner motivation and performance.

VR training has also been implemented in pharmaceutical and biotechnology industries, specifically for training laboratory staff and employees working with customers. Baceviciute et al. [11] conducted a study across Brazil, Denmark and the USA, evaluating VR and video-based training for customer-facing employees in the biotech sector. The study involved 86 participants, evenly split between VR and video-based training groups. The results demonstrated that the VR-trained employees had 41% higher score in knowledge retention assessments and 61% higher in spatial understanding in comparison to the employees who received video-based training. Another notable benefit from the study was that complex biomedical processes could be simulated in VR, which allowed employees to see molecular interactions and laboratory processes in 3D. This resulted in better comprehension and practice of scientific concepts, especially for new employees without previous lab experience.

In the industrial and manufacturing sectors, Singhaphandu and Pannakkong [14] reviewed multiple VR training applications targeting tasks such as assembly line work, machine operations, and emergency preparedness. Based on findings from various case studies, they reported that VR training has been associated with up to 38% improvement in procedural accuracy, a 45% reduction in training time for complex machine tasks, and a 25% decrease in emergency response time. Additionally, several cases showed that VR-trained participants demonstrated higher engagement and retention in some cases up to 50% greater than those trained through traditional methods.

A more comprehensive meta-analysis by Abich et al. [12] synthesized findings from 21 peer-reviewed studies across industries such as defense, healthcare, aviation, and manufacturing. The analysis grouped VR training benefits into three categories: psychomotor performance, knowledge acquisition, and spatial ability. VR consistently enhanced fine motor skills and accuracy in tasks like welding and surgery, improved long-term knowledge retention by 30–50%, and strengthened users' spatial reasoning and situational awareness. Furthermore, 90% of the studies in the review reported increased learner engagement and motivation, 82% showed higher skill retention compared to traditional methods, and 67% documented reductions in human error during real-world task execution.

Despite these promising results, the implementation of VR training has several limitations. Checa and Bustillo [10] reported that approximately 27% of users experienced motion sickness during immersive VR sessions, which negatively impacted their learning experience. Abich et al. [12] also noted that the high upfront costs of VR headsets, tracking systems, and custom content development can pose a significant barrier, particularly for small- to mid-sized organizations.

To bring together the reviewed studies, two summary tables are provided below. Table 3.1 lists each paper's basic details: authors, industry, study type and training type. Table 3.2 then shows the main results, which are key metrics and findings.

| Author(s) | Industry | Design | Training Type |
|---|---|---|---|
| Osti et al. [9] | Construction | Experimental | Assembly (wood wall framing) |
| Baceviciute et al. [11] | Biotech (corporate training) | Experimental (RCT) | Customer product onboarding |
| Singhaphandu & Pannakkong [14] | Cross-industry industrial training (VR/AR/e-learning) | Review | Safety and procedural training contexts |
| Abich et al. [12] | Multi-sector (e.g., surgery, aviation, education) | Systematic Review | Psychomotor, spatial, and conceptual learning (factual learning mixed) |
| Checa & Bustillo [10] | Multi-domain (education, military, health) | Systematic Review of 135 VR serious games | Conceptual and skills learning via serious games |

**Table 3.1:** Summary of context, design, and training type in the reviewed VR Studies.

| Author(s) | Key Metrics | Summary of Findings |
|---|---|---|
| Osti et al. [9] | Task completion time, error rate, user engagement | VR learners completed tasks faster, made fewer errors, and reported higher engagement than video-based trainees. |
| Baceviciute et al. [11] | Conceptual/spatial knowledge, self-efficacy, enjoyment, perceived learning, factual recall | VR significantly improved conceptual and spatial learning, self-efficacy, and enjoyment (d = 0.41–1.74), but not factual recall. |
| Singhaphandu & Pannakkong [14] | Training technologies, application domains, design challenges | Comprehensive review of VR/AR/MR in industrial contexts; highlighted usability, cost, and adoption barriers across sectors. |
| Abich et al. [12] | Psychomotor, spatial, conceptual learning; motivation; factual recall | VR outperformed traditional methods in most domains except factual recall, which showed mixed results across studies. |
| Checa & Bustillo [10] | Game mechanics, immersion, skill/conceptual training, evaluation strategies | VR serious games improved engagement and learning, but evaluations often lacked standardized protocols or validated assessment tools. |

**Table 3.2:** Key metrics and key outcomes in the reviewed VR studies.

## 3.3 Research Gap and Study Contribution

The case studies reviewed in this chapter employed a combination of objective and subjective methods to assess training effectiveness. Most commonly, researchers compared VR training outcomes to traditional methods such as video tutorials or classroom instruction. For example, Osti et al. and Baceviciute et al. used controlled experiments, measuring factors such as task completion time, error rates, and knowledge retention to assess immediate training outcomes. Similarly, Singhaphandu and Pannakkong reported improvements in procedural accuracy and engagement through case study observations in industrial settings. While these studies provide valuable insights into the short-term benefits of VR training, their evaluation strategies are often limited in scope.

Although performance metrics and participant surveys were used, few incorporated expert review, iterative refinement of system design, or the use of structured requirement specifications. Important aspects such as cognitive load and system usability were rarely assessed using standardized tools. Long-term impacts, such as skill retention over time or real-world applicability of acquired skills, were generally not explored in depth.

While existing research demonstrates the effectiveness of VR training across multiple industries, there is a noticeable gap in studies focusing specifically on battery manufacturing safety. Moreover, many of the reviewed studies primarily evaluate VR training from a general skill acquisition perspective, with limited focus on integrating HCI principles and iterative, user-centered development methodologies. This thesis advances the field by applying an evaluation approach that combines user testing, stakeholder feedback, and iterative design improvements in a high-risk industrial context. By using the battery industry as a case study, this research provides validated insights and a transferable framework for developing effective VR training solutions in other hazardous environments.

Furthermore, by focusing on battery manufacturing as the use case, this thesis addresses a critical gap in both industrial application and academic research. The study contributes a validated requirement specification that addresses both general VR training needs and battery-specific safety considerations, directly supporting RQ1. In parallel, the study integrates HCI principles throughout the design and evaluation process, aligning with RQ2 to enhance usability, engagement, and learning effectiveness. By embedding these practices, this study not only contributes to improving VR training within battery production but also provides a reusable design model for similar safety-critical contexts.

# 4

# Methods

This chapter outlines the research design and methodology used to explore the role of VR training in manufacturing workforce education, with a focus on evaluating and improving VR-based battery safety training at BCG. The research followed a structured four-phase approach.

Phase 1 consisted of a pre-study, including a literature search, initial user testing and data collection through surveys and group interviews. Phase 2 involved analyzing the results of the pre-study of the existing VR training system, resulting in a requirement specification for VR-based safety training, both general and use-case specific. Based on these requirements, Phase 3 focused on implementing a new training scenario including the suggested improvements. Finally, Phase 4 evaluated the new VR scenario through additional user testing and validation. The research design combined both qualitative (observations, interviews) and quantitative (survey) data to ensure an extensive evaluation.

Two primary frameworks were used for guiding and designing the methodology. The first one was the ABC framework, which categorizes research approaches based on Actors (A), Behavior (B), and Context (C) [52]. The ABC framework provides a clear and practical way to design and describe empirical research. It facilitates a clear understanding of the necessary trade-offs between studying realistic settings, measuring behavior accurately, and making results generalizable. Since this study focuses on evaluating a VR training system in a real-world industrial context, the ABC framework was well suited for guiding the research design.

The ABC framework defines eight different research strategies that can be used depending on what type of research it is. This study primarily followed the Field Experiment strategy. In a Field Experiment, an intervention is introduced in a real-world environment and its effects are measured while maintaining the setting as realistic as possible. In this study, the intervention was the newly developed VR battery safety training scenario, designed to simulate real industrial tasks and hazardous situations. This was also the independent variable in the evaluation phase (Phase 4). The dependent variables were usability and perceived learning outcomes, measured through a usability survey, cognitive workload, evaluated with the NASA Task Load Index (NASA-TLX) survey and task performance in a real-world battery handling scenario. These measures were selected to provide a complete understanding of the effectiveness and user experience of the VR training intervention.

The second framework was a hybrid approach combining elements from Agile development and Rapid Prototyping, as described in Section 2.2.2. The iterative process of user testing and feedback collection was based on the Agile methodology. Figure 4.1 illustrates an overview of the methodology.

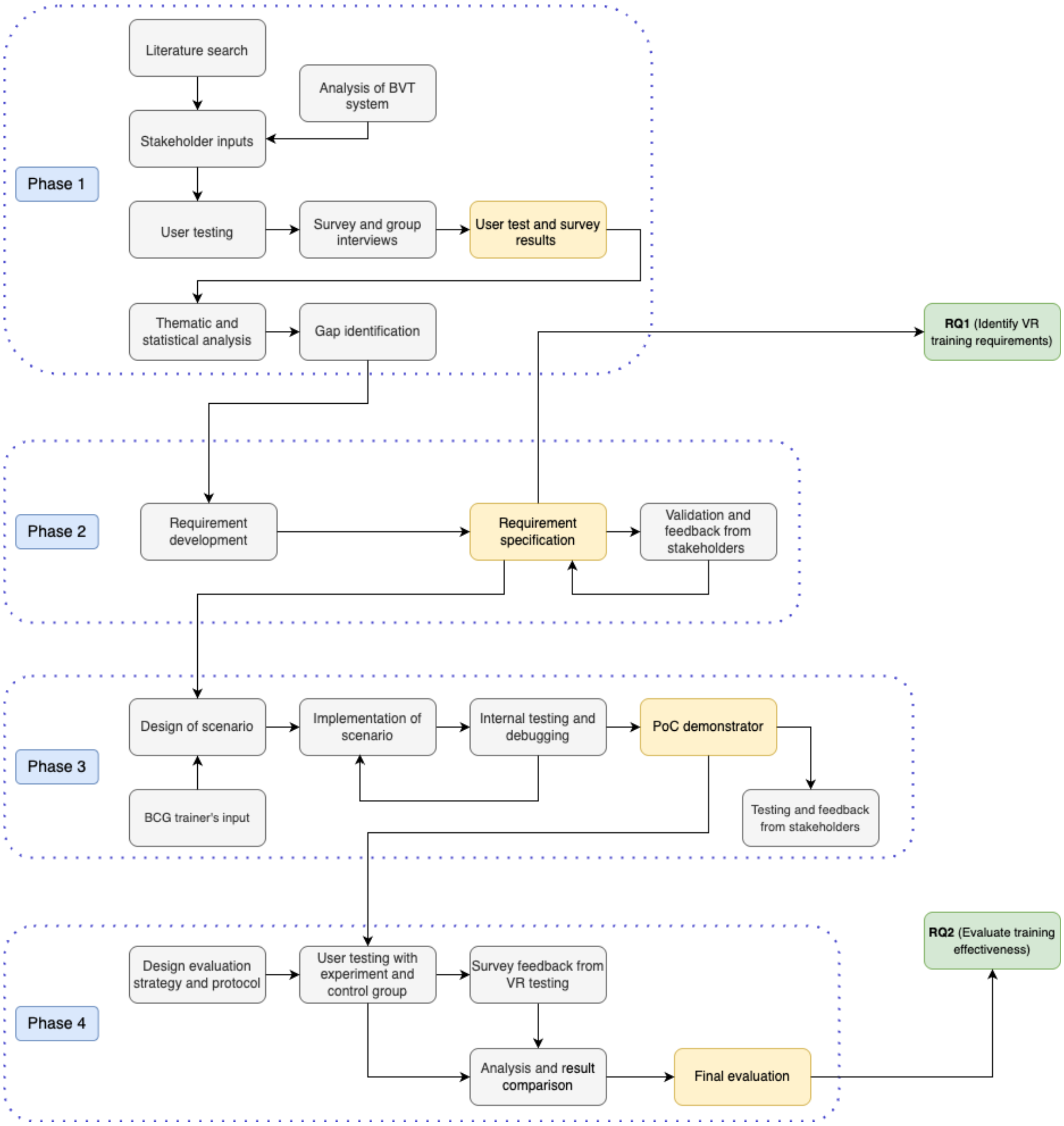



**Figure 4.1:** An overview of the research process across four phases, highlighting main activities (gray), artifacts (yellow), and their relation to the research questions (green).

## 4.1 Phase 1: Pre-study

In the first Phase, the goal was to set a solid foundation for evaluation and enhancement of the VR system. This phase included a literature search, initial user testing, and data collection through surveys and group interviews. It also included a detailed description of the current system and stakeholders' inputs.

### 4.1.1 Literature Search

The very first part of the pre-study was to conduct a literature search, where existing research on VR-based workforce training and HCI principles relevant to VR training were mainly investigated. The publication libraries Google Scholar [53] and IEEE Xplore [54] were used to conduct the literature search. Search queries included combinations of keywords such as "VR industrial training," "battery manufacturing safety training", "human-computer interaction in VR", and "VR history". Inclusion criteria focused on peer-reviewed articles with no specific publication year. The papers were selected based on relevance after reviewing titles and abstracts, and, where necessary, full texts. This part of the pre-study aimed to find best practices and common challenges in VR training. The literature search also helped in creating the basis for the requirement specification, by providing already established usability principles.

The literature search resulted in around 80 initial papers. After titles and abstracts screening, 42 were chosen for full-text review. Of these, 28 were found to be highly relevant and used for detailed analysis. Additionally, snowballing was used to find 7 more relevant sources.

### 4.1.2 Description of the Current System

To fully understand the current system, the Battery Virtual Training (BVT) platform was tested extensively and analyzed. During the analysis of the full VR training flow from beginning to end, key features, user actions, and training content were documented systematically. User stories were used to show how the learning was meant to progress in order to capture both the technical and pedagogical aspects of the system. This section describes the current BVT system in detail, outlining its intended use, pedagogical design, and scenario structure.

The BVT system is a VR-based safety training platform and was created together with GTC to be used in BCG. The BVT system trains workers to handle battery assembly and respond to hazards in an industrial setting. The system runs on the Meta Quest 3, see Figure 4.2. The training focuses on the battery module and pack integration stage, as previously mentioned in Section 2.7. In this stage, workers combine battery cells into modules and assemble them into full battery packs.

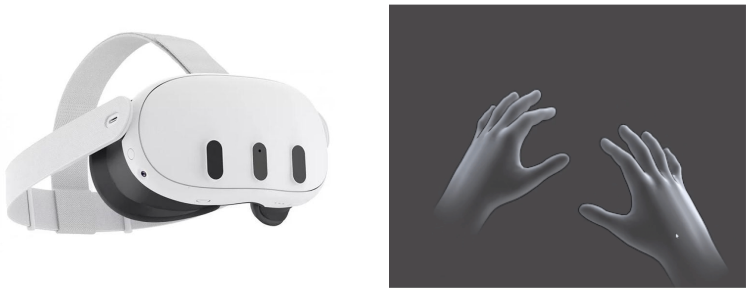


**Figure 4.2:** VR headset (left) and hand-tracking interface (right). The system is used for battery safety training, taken from [55].

Since this process involves live battery cells, electrical safety is a major concern. The goal of the system is to provide hands-on training to help workers learn safe handling techniques, emergency procedures, and proper assembly steps. The training also takes place in a realistic, safe and controlled virtual environment. The BVT system features hand-tracking technology, allowing users to interact with objects naturally by grabbing, pinching, and pressing without using controllers, see Figure 4.2. The users can navigate through the environment simply by walking, but for extreme movement, they can point with their index finger and thumb and perform a pinching gesture to teleport.

The BVT system consists of three main scenarios: "Contact with Battery", "Shorted Battery" and "Smoking Battery". In addition to these scenarios, there is a lobby scene where the user begins and returns to between each scenario. There is also a final exam stage, where users are expected to respond and follow safety measures based on their previous experiences in the safety scenarios while completing the full process of assembling a battery pack. Figure 4.3 presents an overview of the BVT training process, illustrating the system's workflow through a structured flowchart. Throughout the system's main scenarios, the user receives guidance through floating pop-up-like instructions to provoke a hazard situation. This also helps guide the user on what steps to take to handle it. The system also guides the user on what instrument to pick up by highlighting it with a yellow color.

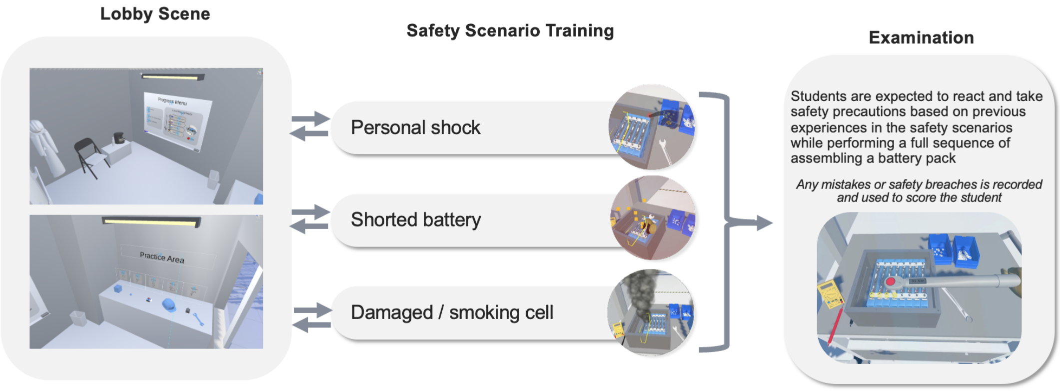
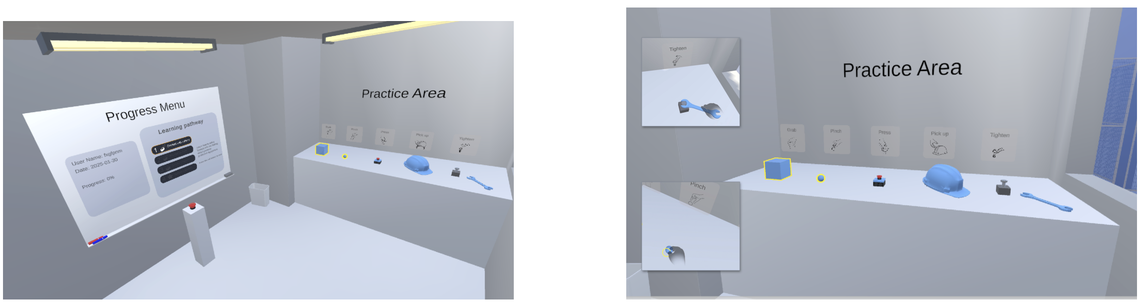



**Figure 4.3:** Overview of the training process in the BVT system, taken from [55].

### Scenario 0 – The Lobby Scene

The user starts at the lobby scene, where the practice area and a whiteboard are presented, see Figure 4.4a. In the practice area, the user gets to learn the basics of interacting with the VR environment. This includes testing different actions such as grabbing, pressing, pinching, and tightening objects, see Figure 4.4b. This step ensures that users become comfortable with the VR controls before engaging in critical safety scenarios. The whiteboard represents a progress menu for the user. It shows a list of all the scenarios and which scenario the user has done. It also shows a percentage based on how far the user is in the progress. Finally, when completing the final exam, the whiteboard will show how long it took the user to complete and how many incidents there have been. All this data are gathered and saved anonymously. In front of the whiteboard, there is a red button that the user needs to press to start each scenario.




**(a)** The lobby scene.

**(b)** The practice area.

**Figure 4.4:** The lobby scene setup in the BVT system, taken from [55].

### Scenario 1 – Contact with Battery

In Scenario 1, the user is introduced to the dangers of improper battery handling and the correct emergency responses. In Step 1, the user is encouraged to touch the battery with their hand and get a burn injury, showing why direct contact with a battery is dangerous. The system guides the user to immediately wash their hand under running water to cool the burn and then call for assistance to report the incident and seek medical support. The next step is to encourage the user to touch the battery again, but with a yellow ring and a clock. This simulates a real-life accident where metal objects can cause sparks and dangerous debris. The user must immediately wash their eyes to remove any tiny particles and call emergency services to get help. Finally, the user must repeat the same steps, but this time while wearing the correct safety gear. They need to put on insulated gloves to protect their hands from burns and safety goggles to shield their eyes from flying debris. By first showing what happens without protection and then introducing safety gear, the scenario helps users build good habits and understand how to stay safe when working with batteries.

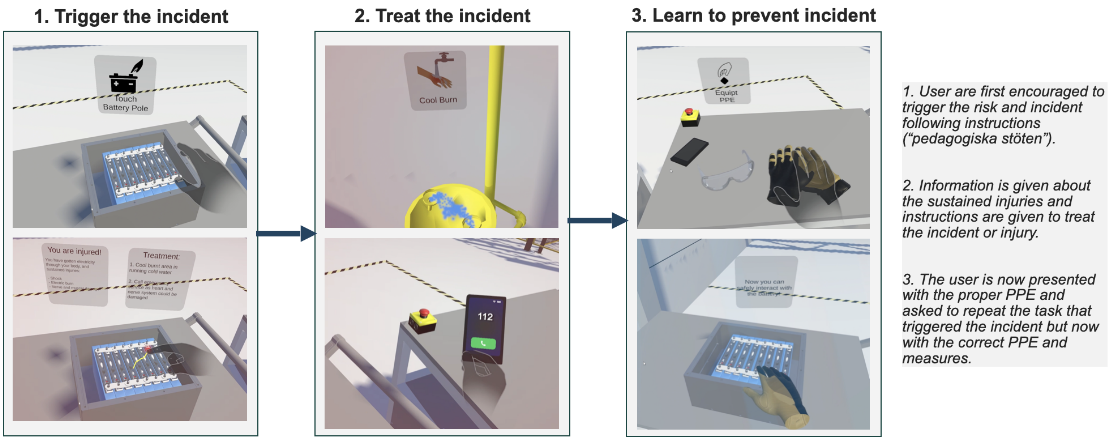



**Figure 4.5:** A general breakdown of Scenario 1 in the BVT system, demonstrating the risks of improper battery handling and the correct safety procedures [55].

### Scenario 2 - Shorted Battery

Scenario 2 focuses on the dangers of a short-circuited battery and how to respond correctly to prevent accidents. The system begins with instructing the user to drop a metal object onto the battery pack, potentially causing a short circuit. The user must quickly measure the battery's temperature using a thermometer to check for any signs of thermal runaway. A thermal runaway is a dangerous condition where the battery rapidly overheats and may catch fire. If the temperature starts rising, the user must evacuate immediately to avoid exposure to heat or fire. If the temperature remains stable, the user should call for assistance instead of trying to handle the situation alone.

### Scenario 3 – Smoking Battery

The final scenario focuses on what happens if you drop a battery cell and the response to a smoking LIB. In the first step, the user is instructed to pick up a battery cell highlighted with yellow and let its poles to touch the battery module placed on the table. This action simulates a short circuit, causing the battery to emit smoke. At this moment, the correct action is to evacuate immediately and follow emergency protocols. In the second step, the user is informed to drop a battery cell on the floor, which could cause internal damage and chemical spill. After this, the user must use a thermometer to measure the battery's temperature. If the temperature rises, it indicates a thermal runaway, which requires immediate evacuation. If the temperature remains stable, the user should call for assistance to ensure the situation is handled safely.

#### Final Exam – Battery Assembly

In the final exam, the user is expected to users apply everything they have learned throughout the previous scenarios. This exam evaluates their ability to assemble a battery pack safely and efficiently, following the correct procedures to minimize safety risks. The system places a ring and a watch on the user's arm and expects the user to realize it and remove them. Additionally, as shown in Figure 4.6, essential safety equipment is placed on a separate workstation. The user must independently identify and use these items without receiving direct instructions, relying instead on the knowledge gained in earlier training. The system highlights the necessary equipment, battery cells, busbars, or nuts in yellow, guiding the user on what to use next. At the end of the exam, the system provides performance feedback, evaluating the user's assembly skills based on total number of safety incidents and the time taken to complete the assembly.

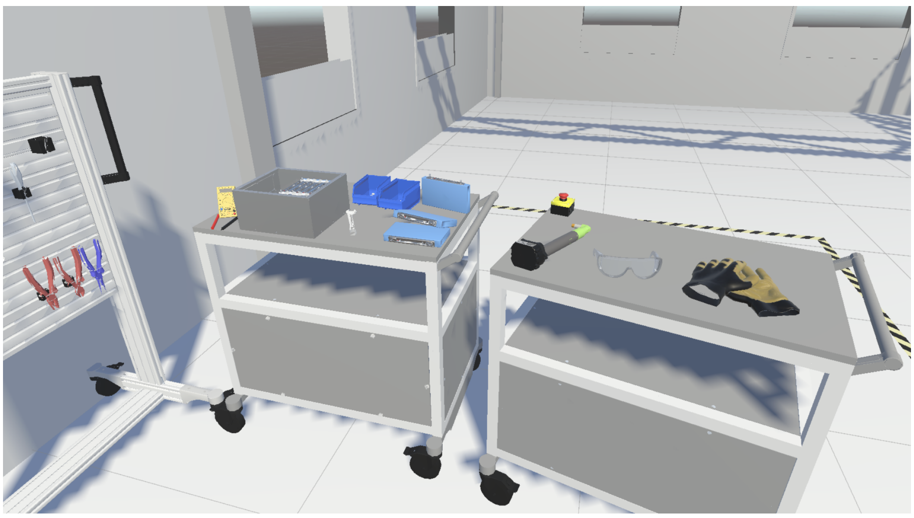


**Figure 4.6:** The workstation setup in the BVT system, taken from [55].

The BVT system provides an immersive and interactive experience, but it is important to evaluate its effectiveness and validate it. The following section presents related work on VR training findings and their results.

### 4.1.3 Initial User Testing

To gain insights into the effectiveness of the BVT system, group sessions of participants testing the system were held. There were six sessions of approximately 2-4 hours, with 5-6 participants in each group, resulting in a total of 78 participants. 48 of the participants worked or had been recently employed to work at a Swedish Battery Assembly Facility, which will be referred to as BAF, and 30 participants were studying quality assurance within battery production at a Higher Vocational Education (HVE) institution. The participants were asked to undergo training in the existing VR system while being observed on their interactions, behavior and encountered issues. The goal of this user testing was to gain feedback on areas such as navigation, ease of use, clarity of instructions and cognitive load. Demographic information was collected to better understand the participants. Table 4.1 shows their age, battery and VR experience, and safety training background. Table 4.2 presents their job roles or study program.

| **Category** | **Subcategory** | **BAF (n = 48)** | **HVE (n = 30)** |
|---|---|---|---|
| **Age** | Under 18 | 1 (2.1%) | 0 (0%) |
| | 18–25 | 14 (29.2%) | 9 (30%) |
| | 26–35 | 23 (47.9%) | 15 (50%) |
| | 36–50 | 6 (12.5%) | 6 (20%) |
| | 50+ | 4 (8.3%) | 0 (0%) |
| **Worked with battery management** | Yes | 13 (27.1%) | 1 (3.3%) |
| | No | 35 (72.9%) | 29 (96.7%) |
| **Training in battery safety** | Yes | 30 (62.5%) | 11 (36.7%) |
| | No | 18 (37.5%) | 19 (63.3%) |
| **Previous VR experience** | Yes | 23 (47.9%) | 16 (53.3%) |
| | No | 25 (52.1%) | 14 (46.7%) |

**Table 4.1:** Participant demographics by group, including age and background experience.

| **Group** | **Roles** | **Participants** |
|---|---|---|
| BAF | Production / Operators | 17 (35.4%) |
| | Logistics | 4 (8.3%) |
| | Maintenance | 5 (10.4%) |
| | Quality / Analysis | 8 (16.7%) |
| | Team leaders | 3 (6.3%) |
| | Specialists / Technicians | 5 (10.4%) |
| | Other / Non-specified roles | 6 (12.5%) |
| HVE | Quality assurance / Battery production | 30 (100.0%) |

**Table 4.2:** Participant demographics by group, including roles and study program.

### Survey and Group Interviews

In the same user testing sessions, after the user testing was completed, the participants were asked to answer a survey and participate in a semi-structured group interview. These were conducted to gain deeper insights into user experiences and perspectives on VR training, providing both quantitative and qualitative data. The interview questions are listed in Appendix A.1 and the survey questions in Appendix A.2.

The survey included 20 questions, divided into four main categories: background (demographic) information, usability and user experience, effectiveness of learning and knowledge retention, and technical and design improvements. The survey was distributed to all VR training participants and responses were collected anonymously. Quantitative data was collected using Likert-scale questions (1-5 ratings), while qualitative data was gathered through open-ended feedback fields. The demographic data included participants' age, prior experience with battery management and safety training and familiarity with VR technology.

After the survey, semi-structured group interviews were conducted with the participants. This allowed for more open ended questions, which meant that the participants felt more free to speak their mind. A total of eight group interviews were held, which were all conducted during the first three test sessions. The interviews aimed to provide deeper and nuanced insights into participants' experiences, elaborate on their survey answers and provide additional feedback. Each group interview lasted approximately 10–20 minutes and discussions were audio-recorded, then transcribed for further analysis. The interview questions focused on overall experience, the effectiveness of VR in learning battery safety procedures, realism of the training compared to traditional methods, usability and interaction issues and suggestions for improvements and system functionality.

All participants were informed about the study purpose and the data handling process before participation. Participation was voluntary, and consent was obtained for both survey responses and audio-recorded group interviews. All collected data were anonymised and no personally identifiable information was collected or stored.

The data from the survey and group interviews were analyzed using both quantitative and qualitative methods. The aim was to identify trends and improvement areas for the VR training system. Since the survey consisted of both qualitative (open-ended questions) and qualitative responses (Likert-scale questions), the analysis approach was different depending on the type of response. The survey responses were compiled into CSV files and the Likert-scale responses were analyzed and plotted using Python. Descriptive statistics were calculated, such as mean and median for identification of the average user perception and outliers. Standard deviations were also calculated to determine variability in responses. This follows the empirical research framework proposed by Kitchenham et al. [56]. A comparative analysis was conducted for determining potential differences in the two participant groups.

The interviews and open-ended responses from the survey were analyzed manually using thematic analysis, following Braun and Clarke's [57] six-phase framework. This included familiarization with the data, generating initial codes, searching for themes, reviewing and refining themes and producing the final results. All group interviews were transcribed and translated from Swedish to English and minor grammatical adjustments were made to improve clarity without altering meaning. The next step included a coding approach to identify both explicit and underlying meanings. Initial codes were applied to relevant text segments. In total, the coded instances found were grouped into 12 recurring code categories for the open-ended survey responses and 15 from the group interview, including usability, instruction clarity and equipment handling. From these codes, eight final themes were developed based on conceptual similarity and relevance to training outcomes from the survey. Similarly, ten themes were finalized from the group interviews. A detailed summary of the codes and themes elicited is included in Appendix A.5.1 and A.5.2.

### NASA-TLX Survey

An additional testing of the BVT system was conducted with six participants. In this session, after completing the testing, the participants responded to the NASA-TLX survey, which is a tool for assessing cognitive workload. It captures workload across six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration [58]. Participants rated each dimension on a scale from 0 to 100, where higher values indicate greater perceived workload [59]. Figure 4.7 shows the definition of each dimension and the rating endpoints.

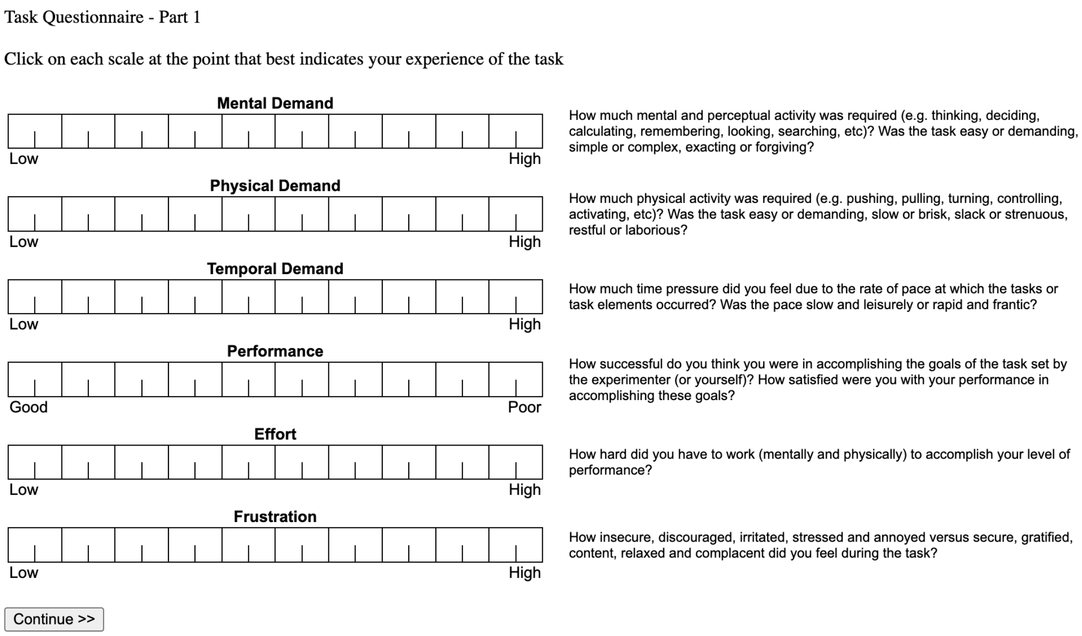



**Figure 4.7:** Overview of the dimensions, their definition and the endpoints. Taken from [60].

The later part of the NASA-TLX survey accounts for individual differences in perceived workload structure, by performing pairwise comparisons between each dimension to determine its relative importance to the task. See Appendix A.11 for example questions of this part of the survey. These weights reflect how strongly each workload component contributed to the participant's perceived workload [58]. Weighted scores are then added and divided by 15 to get a final workload score between 0 and 100 [59]. The survey was administered using an online version that automatically calculated the workload scores based on the provided answers [60]. The purpose of this data collection was to create a benchmark for evaluating results in Phase 4.

### 4.1.4 Stakeholder Inputs

To understand what is important for designing a VR training system that follows real-world safety training standards, key industry stakeholders were consulted. Specifically, insights were gathered through an in-depth interview with two training coordinators at BCG. The interview questions are listed in Appendix A.3. Their contributions provided valuable context on the structure and challenges of traditional training methods used at the centre and their expectations for new training technologies such as VR. Topics covered included current safety procedures, learning goals, training delivery formats, and their perspectives on VR as a training complement.

In parallel, ongoing discussions were held with Henrik Söderlund [61], the developer of the existing BVT system. He provided insights into the original design rationale of the system, including its pedagogical flow, scenario selection, and technical considerations. The insights gathered from these stakeholders were incorporated into the requirement specification and the design of the improved VR scenario.

## 4.2 Phase 2: Development of Requirement Specification

In Phase 2, a requirement specification was developed based on user feedback, literature findings, HCI principles, and interviews with stakeholders to understand industry requirements and training standards. The specification included a foundation for designing VR safety training systems for industries, including a section specifically for battery safety training.

A requirements specification framework was followed to structure software requirements [44] and aligned with the IEEE 830-1998 Software Requirements Specification Standard [45]. It was then uploaded to the scientific open access platform Zenodo.org [62] for spreading the results to be further used by industry practitioners and VR training system developers.

The requirement specification was categorized into the following parts:

- **Introduction:** defined the purpose, scope intended application of the VR training system. This was determined through stakeholder inputs.

- **System requirements:** defined general structure and technical requirements of the VR training system, including hardware and software requirements. These requirements were gathered through industry and stakeholder inputs and the literature search.

- **Data requirements:** described the types of data collected, stored, and processed within the system, for instance performance metrics.

- **User requirements:** outlined user needs and considerations, and usability factors to ensure an engaging learning experience. These requirements were based on user testing and HCI principles, as well as the survey and group interview analysis.

- **Functional requirements:** specified core system features and interactions needed to facilitate learning. These requirements were established from the three primary data sources used in this study: stakeholder input, user feedback and literature search.

- **Non-Functional requirements:** specified factors such as system performance, usability and scalability, ensuring that the VR training is realistic and effective. The data outlining this part of the requirement specification is the same as for the functional requirements.

To systematically translate user insights and research findings into a formal requirement specification, a structured multi-step process was followed. An overview of this structured process, from data collection to the finalized requirement specification, is illustrated in Figure 4.8. The process included the following steps:

1. **Data collection:** Survey responses, group interview transcripts, user observations, and stakeholder interviews were collected during Phase 1.

2. **Analysis:** Recurring problems, user needs, and improvement suggestions were grouped into major themes such as usability, realism, task clarity, and technical reliability.

3. **Mapping to categories:** Each theme was mapped into one or more requirement categories: Functional Requirements, Battery-Specific Functional Requirements, Non-Functional Requirements, System Requirements, User Requirements, and Data Requirements.

4. **Formulation of requirements:** For each identified theme or user need, a concrete requirement was formulated. Requirements were structured according to the IEEE 830-1998 standard for software requirements specifications [45].

5. **Prioritization:** Each requirement was assigned a priority level (high, medium or low) based on its criticality for user safety, impact on learning effectiveness

and relevance for usability and user satisfaction.

6. **Validation with stakeholders:** The requirements were reviewed with trainers from BCG to verify the completeness, clarity, and practical relevance of the specification before starting the implementation of Phase 3. The requirement specification was validated category by category and feedback was documented systematically during the session. Minor clarifications were suggested, such as more explicit wording for emergency scenario procedures, enhanced emphasis on contamination risk handling and some changes in the prioritization.

7. **Finalization:** After stakeholder validation, the finalized requirement specification was completed, containing 61 categorized requirements.

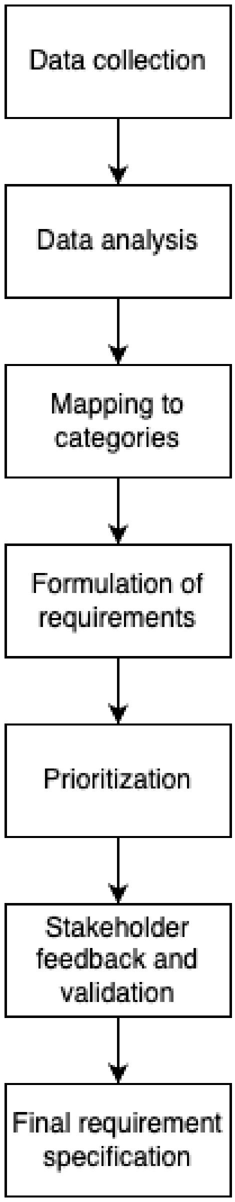



**Figure 4.8:** Overview of the process conducted in Phase 2, from data collection to validated requirement specification.

## 4.3 Phase 3: Development of a PoC Scenario

To address the gaps found in Phase 1, a new scenario was created as a PoC for the VR safety training system. The goal was to build a short and focused training experience that showed how the missing requirements could be met with a focus on the raw to cell process in battery production. Figure 4.9 provides an overview of the workflow followed during Phase 3.

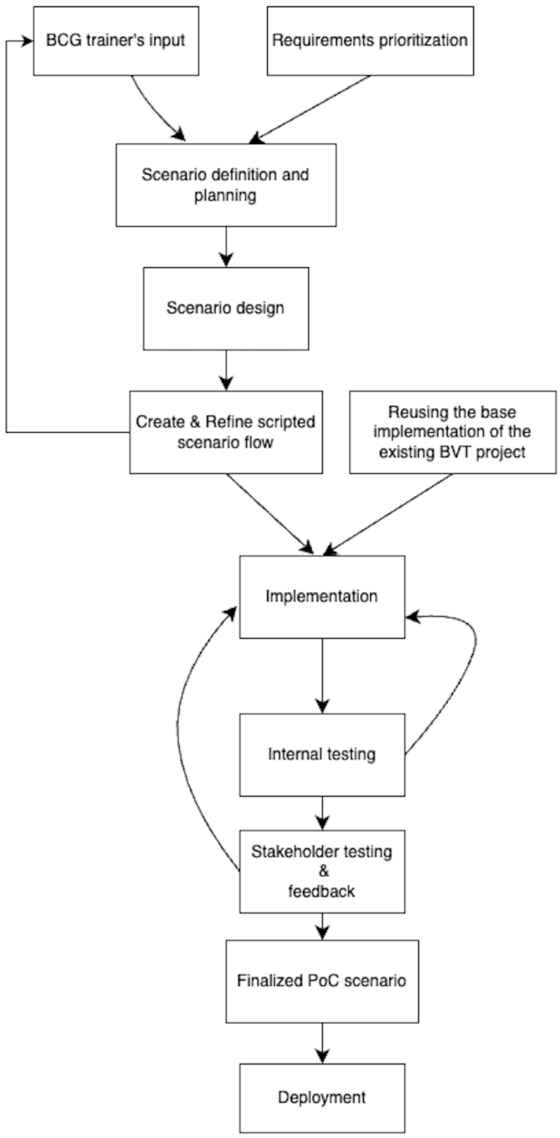



**Figure 4.9:** Overview of the process conducted in Phase 3, from scenario planning to final deployment of the PoC.

The first step was to brainstorm scenario ideas with stakeholders from BCG, to choose the best training event for the PoC. A scenario involving a dropped battery that spills and release colorless gases from its electrolyte was chosen, including an evacuation procedure. These gases are flammable and toxic, requiring the user to evacuate while using protective equipment. As mentioned in Section 2.7, one of these gases is hydrogen fluoride. Even a small amount can cause serious damage to the lungs, eyes, and skin. It can also be deadly if inhaled, which is a good example of a hazardous situation that is life threatening to simulate in real life. Table 2.2 summarizes the hazard classifications for hydrogen fluoride based on international GHS standards [49]. Once the scenario idea was confirmed, the development process begun.

### 4.3.1 Development Process

The implementation followed an iterative and user-focused development process. It started by closely following the Requirement Specification from Phase 2. The main focus was to meet as many high- and medium-priority Functional Requirements as possible, both general and battery-specific. High- and medium-priority User Requirements were also prioritized to improve the user experience and make sure the system included the key features needed for effective safety training. Although the focus was mainly on Functional and User Requirements, some important Non-Functional and System Requirements were also explored. Requirements involving multi-phase progression, branching narrative structures, or final examination scenarios were not prioritized during the PoC implementation due to the limited scope defined for this development phase.

As part of early development planning, a scripted scenario flow was created to describe the expected sequence of user actions and system responses. This script served as a reference throughout the implementation and was refined iteratively based on stakeholder feedback. A draft was reviewed by a BCG trainer to ensure it reflected realistic procedures and met training expectations. The goal was to ensure that the PoC aligned with both the requirement specification and the practical demands of the training environment. Reference images of reused components such as the battery model and a 3D SketchUp object of the face mask (respirator) were also shared to confirm visual and functional suitability.

The new scenario was built by reusing the base implementation of the existing BVT project. This approach allowed the development team to reuse essential components while preserving the stability of the original system. Unity version 2022.3.22f1 was used during development to maintain compatibility with the current VR system architecture. The Meta XR All-in-One SDK (Oculus SDK) [63] was used to provide VR functionality, including essential support for hand tracking, as specified in the requirement specification. Real-time testing and iteration were possible by using Meta Quest Link, allowing developers to test scenario components directly within the VR environment as they were implemented. Meta Quest 3 headsets were used for final deployment and internal testing. Object assets were sourced from SketchUp 3D Warehouse [64] to accelerate the development of the training environment. Voice instructions were generated using the NaturalReaders online platform [65], providing audio cues during hazard escalation phases. Technical guidance and setup support were provided by supervisor Henrik Söderlund [61] to ensure that development practices remained aligned with the current system standards.

To ensure development efficiency and design coherence, the implementation plan included the reuse and adaptation of existing assets from the current VR system. The lobby scene was selected as a foundation for constructing the clean room environment, with modifications planned to remove non-essential 3D models and maintain a sterile appearance. Core UI components such as the canvas structure and reset/restart buttons were intended to be reused to support consistent interaction behavior. In addition, key visual elements, including the battery model, smoke, and leakage effects were to be integrated and adjusted to support the hazard escalation sequence outlined for the new scenario.

Two scripts in C# were developed to control the new scenario's specific flow. A new script, `CellDrop.cs`, was written from scratch, replacing the previous version. It manages the full sequence of events from the start of the scenario through hazard escalation to evacuation, see Appendix A.7.1. `Head.cs` was created to detect when the user's head collided with the exit zone, triggering the evaluation function within the scenario logic, see Appendix A.7.2.

Continuous testing was conducted during development to verify each feature and ensure a smooth scenario flow. Key elements such as interaction mechanics, instruction sequences, visual and particle effects, audio cues, UI components and progress tracking were regularly tested and refined. Additionally, privacy considerations were addressed by ensuring that no personal data was stored and all test data remained anonymous. After completing these tests, the final scenario was deployed on Meta Quest 3 for formal evaluation. Both stakeholders from BCG got to test the final version of the PoC themselves and provided feedback, which led to final adjustments.

## 4.4 Phase 4: Evaluation of New Scenario

The final phase of the study focused on evaluating the PoC scenario through experimental user testing and feedback. The testing and evaluation design in this phase was inspired by De Giorgio et al. [2], who emphasize the importance of using structured experimental approaches when validating XR-based training systems. Following their recommendations, Phase 4 was designed with both quantitative and qualitative methods to assess the PoC's effectiveness. Objective task performance outcomes between VR-trained and text-trained participants were measured. Survey feedback on training experience, usability, and cognitive workload was also collected, including both quantitative and qualitative data. The goal was to assess whether the enhancements implemented in Phase 3 had successfully addressed the identified issues and improved usability, engagement and learning effectiveness compared to both the BVT system and traditional training.

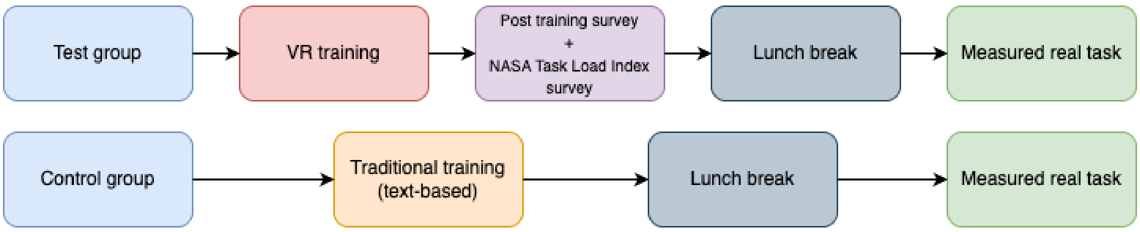



**Figure 4.10:** Experimental design overview for Phase 4.

The quantitative study was conducted as a controlled experiment with two groups: a test group (VR-trained) and a control group (text-trained). Figure 4.10 illustrates the structure and flow of this experimental design. The test group received VR-based training by testing the scenario developed in Phase 3, while the control group received traditional text-based instructions on chemical safety, derived from BCG's training material. 14 participants took part in the physical testing, with 7 participants in each group. At the start of each session, all participants received a 10-minute introduction presenting the research, the purpose of the testing and basic information about how to interact with the VR system. Following the introduction, participants were randomly divided into the two groups.

The first group was the test group, where the participants tested the VR scenario individually. After completing the scenario, the participants completed two surveys: a post-training usability survey on perceived training effectiveness and a NASA Task Load Index (NASA-TLX) survey measuring cognitive load. Each participants' VR testing and survey completion lasted approximately 10–15 minutes. The second group, the control group, received a small booklet covering chemical handling, general lab rules and behaviour and PPE usage. They were instructed to read carefully for 15 minutes in a quiet environment without interruption. After the training phases, all participants took a 30–60 minute break before proceeding to the real measured task.

As in Phase 1, demographic data was gathered through the survey responses. Table 4.3 presents details on the age ranges, prior experience with battery management and safety training, as well as familiarity with VR among the Phase 4 survey respondents.

| Category | Subcategory | Participants (n = 12) |
|---|---|---|
| **Age** | Under 18 | 3 (25%) |
| | 18–25 | 5 (41.7%) |
| | 26–35 | 3 (25%) |
| | 36–50 | 0 (0%) |
| | 50+ | 1 (8.3%) |
| **Worked with battery management** | Yes | 0 (0%) |
| | No | 12 (100%) |
| **Training in battery safety** | Yes | 2 (16.7%) |
| | No | 10 (83.3%) |
| **Previous VR experience** | Yes | 11 (91.7%) |
| | No | 1 (8.3%) |

**Table 4.3:** Phase 4 survey participant demographics.

### 4.4.1 Surveys and Data Analysis

Survey feedback was collected from all 7 VR-trained participants and an additional separate group of five participants who evaluated only the VR training scenario without participating in the real measured task. All participants completed two surveys immediately after completing their respective VR training sessions.

Adopting the same structure as in Phase 1, participants conducted two surveys, the usability survey and the NASA-TLX survey. The usability survey was nearly identical to the survey from Phase 1, as described in Section 4.1.3. For this phase, an additional section with questions about the PoC was added to collect specific feedback, see Appendix A.10.

To evaluate the effectiveness of the PoC, data from this phase was analyzed and compared with results from Phase 1. Multiple types of data were collected and analyzed:

- **Performance metrics:** Measures task completion time, correct use of PPE, battery inspection, hazard recognition and successful evacuation.
- **Survey data:** Ratings from Likert-scale responses on usability and training effectiveness, as well as responses from the NASA-TLX workload survey.
- **Qualitative feedback:** Open-ended survey responses and notes captured during the real measured task.

Quantitative analysis was conducted using descriptive statistics (mean and standard deviation) for Likert-scale survey responses and NASA-TLX scores. Task success rates and performance outcomes were compared between the test (VR-trained) and control (text-trained) groups. Comparisons were also made with Phase 1 VR data to assess improvements in perceived usability, workload and behavioral task performance. Qualitative analysis involved thematic coding of open-ended responses and behavioral observations. Recurring issues, suggestions and positive remarks were identified and grouped into themes.

### 4.4.2 Measured Real Task

The participants individually completed a measured real task, designed to simulate an emergency battery incident response. The task was conducted in a dedicated room, including mirrored key elements from the VR scenario and the text-based instructions studied by the control group. Figure 4.11 shows the real-world setup.

The setup included a table with a mock battery cell, a surgical mask, a chemical gas mask (labeled "respirator" for clarity), gloves and hand sanitizer. Additional items, such as scissors, rulers and a calculator, were also placed on the table for distraction. The exit of the room was labeled as the emergency exit through a printed out emergency exit sign, to create an authentic evacuation environment. Hazard simulations were created through audio cues, including a gas leak recording, a spoken warning and a loud evacuation alarm. These stimuli were used to assess hazard recognition and trigger appropriate responses.

While the gas mask was essential for successful task performance and was explicitly expected to be used, the surgical mask, gloves and hand sanitizer were intentionally included as distraction items. Their purpose was to test participants' ability to correctly identify and prioritize the PPE relevant to the simulated emergency, rather than default to general PPE use. This design choice intended to introduce a decision-making part to the task, simulating the cognitive complexity workers may face in real-world safety situations where not all available tools are contextually appropriate.


**Figure 4.11:** Real-world task setup used in Phase 4.

The task flow included the following steps:

1. Pick up the battery cell, drop it on the floor and pick it up again.
2. Inspect the battery for visible damage or leakage.
3. Recognize the simulated gas leak upon hearing the audio recordings.
4. Correctly put on the chemical gas mask provided in the room.
5. Respond to the loud evacuation alarm by quickly finding and exiting through the emergency exit.

Each participant's performance was observed and recorded following a standardized framework protocol that recorded errors and response time of the incident handling by all participants. Participants were anonymized using the format TG-Px for the test group and CG-Px for the control group, when noting their behavior. The full protocol with notes is presented in Appendix A.12.

# 5

# Results

This chapter presents the results of the study, following the structure of the four phases included in the methodology. The findings draw on both qualitative and quantitative data sources, including surveys, group interviews, stakeholder feedback and empirical data measures. Phase 1 presents insights from the initial evaluation of the current VR training system, highlighting user experiences and challenges. Phase 2 details the development of the requirement specification, mapping feedback and theoretical insights into structured requirements. Phase 3 outlines the development outcomes of the improved VR scenario, describing the implemented features and anticipated improvements based on the requirement specification. Finally, Phase 4 presents the feedback from the improved training scenario and a comparison between the Phase 1 and Phase 4 results.

## 5.1 Results from Phase 1

This section presents the results of the initial system evaluation conducted through surveys and group interviews. Both quantitative and qualitative data were collected. The quantitative results provide an overview of user ratings on usability and training effectiveness, while the qualitative analysis identifies recurring patterns in user feedback. The NASA-TLX survey results from this phase are presented with the Phase 4 results in Section 5.4.2.

### 5.1.1 Usability Survey Results

The quantitative data from the survey responses were analyzed to evaluate participants' perceptions of the VR training system across multiple dimensions. Participants rated their experiences on a 5-point Likert-scale, and the results were grouped into ten themes: VR Comfort, Interface Intuitiveness, Immersion, Incident Realism, Confidence, Understanding Risks, Safety Recall, Reflection, Instruction Clarity, and Educational Value.

The demographic data show that most participants were between 26-35 years old, with 47.9% of BAF and 50% of HVE respondents falling into this age range. 27.1% of BAF participants had prior experience with battery management, while HVE participants reported 3.3%. Similarly, 62.5% of BAF participants had previously received battery safety training, compared to 36.7% of HVE students. Full demographic breakdowns by role, age and experience are presented in Table 4.1 and 4.2.

The results are presented in a comparative bar chart with error bars representing the standard deviation (SD), as shown in Figure 5.1. The results show generally high levels of satisfaction across both groups, with most themes scoring close to or above 4 on average. HVE students reported slightly higher comfort using VR, while BAF participants reported higher ratings for confidence and realism. Both groups rated the immersion and educational value themes highly, indicating that the training was both engaging and effective. Additionally, the interface intuitiveness rating was the highest among the HVE participants and second highest among the BAF participants. Themes like VR comfort, incident realism and confidence revealed more variation for both groups. Overall, the SD scores were low in many themes, which indicates that participants responded consistently in the groups. The details of these results are summarized in Appendix A.4.1.

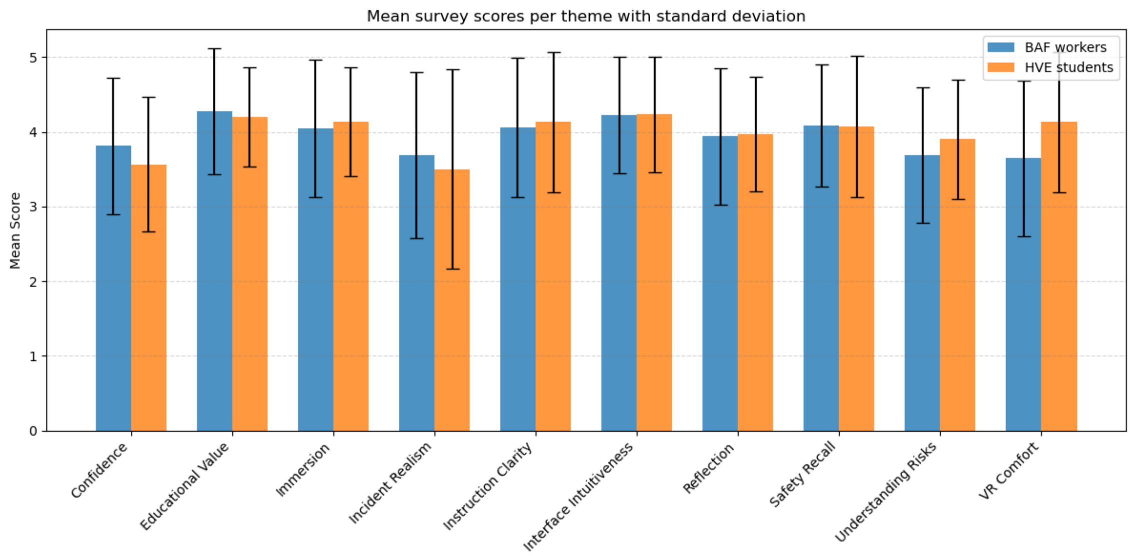



**Figure 5.1:** A bar chart showing the mean scores with SD of quantitative survey responses categorized into themes.

The open-ended survey responses were categorized into eight themes based on thematic similarity. Table 5.1 summarizes the final themes and their descriptions. The detailed coding process, including each individual code, its assigned theme, and frequency of occurrence, is provided in Appendix A.5.1.

Participants frequently commented on the engaging nature of the VR experience. Many highlighted that the interactive elements contributed to a higher level of engagement compared to traditional training methods. The realism of the scenarios was appreciated, especially how the system simulated the consequences of incorrect actions. For example, one survey participant remarked that the training "felt very realistic, which helps one remember the correct procedures better."

From a learning perspective, participants reported that the combination of visual, physical, and auditory cues in the VR environment helped them understand and retain information more effectively. As one participant explained, VR offered a “good way to learn, better than just reading theory.”

However, users also encountered technical challenges that sometimes detracted from the experience. One frequently reported issue was difficulty in interacting with virtual objects. A participant noted that it was "sometimes hard to pick up objects," while another commented that "sometimes it was hard to understand what was expected."

In terms of improvements, participants suggested several enhancements to increase the realism and effectiveness of the training. These included clearer guidance, better feedback from the system, and additional environmental effects. Some participants expressed a desire for more varied scenarios and deeper training modules. Finally, the participants recognized the value of VR training in preparing for real-world work situations, with one participant reflecting that VR was "good to practice before doing it for real."

| **Theme** | **Description** |
|---|---|
| Interaction and usability | Comments related to difficulties with object handling, tool manipulation, or the general control mechanics of the VR system. |
| Technical issues | Reports of bugs, lag, hardware glitches, or unexpected system behavior during the VR experience. |
| Instruction clarity | Suggestions related to how instructions were presented, including clarity, timing, or the need for better task guidance and feedback. |
| Content and scenario improvement | Suggestions for expanding or diversifying the safety scenarios within the training to increase realism or relevance. |
| Realism and immersion | Observations related to how authentic or believable the virtual environment and its physics felt. |
| Physical discomfort | Mentions of mild discomfort, eye strain etc. experienced during training. |
| Collaboration and multi-user | Requests for features enabling teamwork or parallel participation in training sessions. |
| Overall impression | General feedback reflecting users’ emotional reactions, including positive impressions and perceived usefulness. |

**Table 5.1:** Final themes and descriptions derived from open-ended survey responses.

### 5.1.2 Findings from the Group Interviews

The group interviews with participants from BAF revealed ten themes regarding the VR training system. Table 5.2 presents an overview of the final themes and their corresponding descriptions. The full list of individual codes, along with how frequently they appeared and how they map to each theme, is available in Appendix A.5.2.

Many participants described the VR experience as immersive and engaging. They appreciated the opportunity to experience scenarios that would be dangerous in real life, highlighting how the realism of simulated consequences improved their learning. For instance, participants explained that it was "fun to see what happens when you make mistakes that we absolutely must not make in practice," adding that these experiences made the risks more appreciable. Another participant expressed that they "actually jumped" during a simulated electric shock, emphasizing the sense of realism and immersion.

The effectiveness of VR in reinforcing safety protocols and procedures was a frequently mentioned benefit. Participants felt that combining theory with practice through VR enabled better learning outcomes. One participant commented that they "learn more when actually doing something" rather than simply reading about procedures. Other participants highlighted that the experiential nature of VR helped to establish a form of muscle memory, improving the likelihood of recalling critical actions in real scenarios. One participant summarized this by saying that "you combine theory with practice. This is a very good example."

Participants appreciated how the scenarios were structured to progressively build on their knowledge. However, participants also reported technical challenges. Issues such as system lag, difficulty in picking up objects, and unclear feedback from the VR environment sometimes disrupted the training flow. Regarding improvements, participants proposed several ideas to enhance the training experience. They suggested adding more environmental realism, such as factory sounds and visual distractions, to better mimic actual working conditions. One participant noted that "it would have been good if it was more like a factory environment, with some background noise, like presses."

Additionally, some participants expressed interest in multi-user scenarios and differentiated roles within the simulations. One participant remarked, "if you're in the simulation together, maybe with different roles, it would be even more realistic."

Some participants recommended more impactful incident effects, like stronger haptic feedback and louder sounds. They acknowledged the relevance of VR training to their work in the real-world. They saw value in using VR to simulate hazardous situations in a safe environment and suggested expanding the system to cover more specific tasks encountered in their daily operations. One participant commented that they “see that there’s definitely a place for this in our operations.”

Overall, participants agreed that VR training would be highly valuable for onboarding new employees, especially for safely exposing them to hazardous environments.

As one participant put it, "it's great to have gone through this before handling a live battery."

| Theme | Description |
|---|---|
| Realism and immersion | Participants described the VR training as realistic or immersive, highlighting that it felt like a real-life situation. |
| Learning by doing | Emphasis on the effectiveness of experiential learning, where doing tasks helped users understand and remember better. |
| Shock and consequence | Participants reacted strongly to surprising or intense elements (e.g., sounds or accidents), reinforcing learning through emotional impact. |
| Instruction clarity | Comments on whether instructions were clear, confusing, missed, or needed improvement for better task execution. |
| Technical issues | Reports of bugs, glitches, tracking problems, or other disruptions in the VR experience. |
| Suggestions for improvement | Concrete ideas or wishes for additional features, alternative scenarios, or clearer functions. |
| Collaboration and multi-user | Comments suggesting that VR training could benefit from multi-user features or team-based scenarios. |
| Retention and memory | Participants felt they could remember information better after performing tasks in VR. |
| Comparison to traditional training | Reflections on how VR compares to classroom-based or theory-heavy training; many found VR more effective or engaging. |
| Physical discomfort | Mentions of dizziness, eye strain, nausea, or other physical symptoms experienced during VR use. |

**Table 5.2:** Themes from group interviews and their descriptions

### 5.1.3 Observations from the User Testing in Phase 1

During the initial user testing, participants expressed general excitement about the VR training system. However, several usability issues were identified. Some users, especially older participants, were hesitant and needed more time to adjust. Many were afraid of walking into walls or furniture, particularly in group sessions involving four to six people. Casting the headset view to an external screen helped trainers guide users more effectively, but only two devices could be cast at once. This limited the support trainers could offer. Immersion was often disrupted when users felt unsure about what to do next. Scene transitions caused confusion, especially when the screen turned black. Some users panicked or got scared until they were told it was only a transition. When the system lagged or took time to load, users became impatient. Several participants did not recognize the restart button on the table, which showed a lack of affordance. Others found it difficult to read overlapping instructions or understand the correct reading order. In one scenario, smoke effects covered the instruction text, making it hard to know how to act. This caused stress and a delay in user response. Participants who wore glasses appreciated that the Meta Quest 3 headset supported glasses comfortably. Although some struggled with object interactions, most valued the ability to use their own hands in VR. No participants reported experiencing significant motion sickness during the sessions.

### 5.1.4 Findings from the Interview with BCG Trainers

The trainers described the current training structure as a combination of theoretical classroom sessions and practical hands-on exercises. Theoretical parts include lectures and printed materials, but they acknowledged that maintaining participant engagement is a challenge. They noted that "it is difficult. We've even considered breaking up theory sessions with energizers to keep participants awake and attentive".

To address this, their pedagogical approach involves alternating between short theory segments and practical exercises. This strategy helps maintain engagement and allows learners to immediately apply theoretical knowledge in practice. As one trainer explained, "we try to weave in practical parts directly after theory, so they get to apply what they've just learned right away". The trainers emphasized that practical exercises not only increase engagement but also encourage participants to ask questions more freely. Compared to the traditional classroom environment, hands-on sessions promote curiosity and open dialogue. One trainer stated that "people become much more active and curious. They start asking questions, sometimes to each other first, which later leads to group discussions".

However, certain safety-critical scenarios are difficult or impossible to replicate in physical training environments. For example, simulating high-risk incidents like thermal runaway or electrical faults is neither safe nor feasible. Here, the trainers identified a clear opportunity for VR to enhance the training program, stating that "especially for the dangerous scenarios. We can talk about them, but to actually experience them in VR is a huge advantage". Furthermore, the trainers described VR as an effective tool to reinforce critical safety knowledge, providing learners with memorable experiences that complement theoretical instructions. The trainers noted that participants enjoyed the VR experience and found it helpful for connecting theoretical concepts with practical risks. They described VR sessions as resembling "almost a teambuilding activity", because of the high engagement and discussion it generated.

The trainers also saw potential in using VR to onboard new employees, providing them with early exposure to hazardous situations in a safe and controlled environment. One trainer commented that "it's great to go through this before handling a live battery". Despite these benefits, they acknowledged limitations in current training, such as the inability to replicate full PPE conditions or prolonged exposure to workplace hazards. They also emphasized that VR should be well-integrated into the broader training flow, complementing rather than replacing physical practice.

## 5.2 Results from Phase 2

Building on the insights gained from the pre-study, survey responses, group interviews, user testing observations and relevant literature found in the literature search, a comprehensive requirement specification was developed for the VR training system. This document serves both as a general framework for VR-based safety training systems and as a tailored specification for battery safety training, addressing the unique risks associated with battery production environments. The complete specification is published as a standalone document and is publicly available via Zenodo.[1]

The quantitative survey data provided clear indications of user perceptions across themes such as VR comfort, interface intuitiveness, realism, and instructional clarity. These responses highlighted both strengths and areas needing improvement in the existing system, particularly in relation to task guidance, accessibility, and realism of simulated scenarios.

Qualitative responses from the survey and group interviews, together with the observations, enriched this understanding by providing detailed descriptions of user experiences. Themes such as realism and engagement, learning and knowledge retention, and applicability to real work consistently emerged. Participants expressed the importance of realistic hazard simulations, clear instructions, and the ability to learn from mistakes in a safe environment.

Expert input from stakeholder interviews further validated these findings, ensuring alignment with industry standards and practical training needs. Additionally, literature insights on VR design, human-computer interaction, and learning theories (such as experiential learning and cognitive load management) informed system goals such as maintaining high immersion while avoiding cognitive overload.

Based on these combined insights, a total of 61 requirements were defined and categorized into functional, non-functional, system, user, and data requirements. Each requirement was prioritized according to its criticality to user safety, system usability, and training effectiveness. High-priority requirements address essential safety and usability features, while medium and low priorities reflect improvements that enhance user experience and training depth. A summary of the requirement categories is presented in Table 5.3. A full cross-mapping of themes from all data sources is included in Appendix A.6.

Each requirement was prioritized as high, medium or low. The prioritization was guided by clear success criteria to ensure consistency and transparency. High-priority requirements were identified as those directly related to safety-critical tasks and essential training objectives, which are necessary for user safety. These are also necessary to ensure the system has the fundamental features to work as intended.

[1] https://doi.org/10.5281/zenodo.15355792

Medium-priority requirements were selected as features that improve usability, engagement, and performance but are not essential for basic operation. Low-priority requirements were classified as supportive or nice-to-have features that can be implemented in later development stages if needed.

The following prioritization thresholds were defined:

- High-priority requirements: 100% fulfillment is mandatory.
- Medium-priority requirements: At least 70–80% fulfillment is recommended.
- Low-priority requirements: Fulfillment is desirable but not required.

Overall, achieving at least 70–80% fulfillment of all requirements was defined as the success criterion for the VR training system to be considered effective for industrial deployment. This target, combined with full implementation of all high-priority requirements, aligns with best practices in software engineering and XR training system evaluation.

As discussed by Kitchenham et al. [56], the establishment of clear success criteria and measurable outcomes is crucial for system evaluation in empirical software engineering research. Similarly, the IEEE 830-1998 standard [45] underlines the importance of prioritisation to ensure that critical features are fully implemented to achieve system acceptability.

| **Category** | **Total** | **High** | **Medium** | **Low** |
|---|---|---|---|---|
| General Functional Requirements (FR) | 13 | 7 | 5 | 1 |
| Battery Production Functional Requirements (FR-B) | 8 | 6 | 2 | 0 |
| Non-Functional Requirements (NFR) | 12 | 6 | 4 | 2 |
| System Requirements (SR) | 11 | 8 | 3 | 0 |
| User Requirements (UR) | 11 | 6 | 4 | 1 |
| Data Requirements (DR) | 6 | 4 | 2 | 0 |
| **Total** | **61** | **37** | **20** | **4** |

**Table 5.3:** Summary of requirements and prioritizations.

## 5.3 Results from Phase 3

In Phase 3, a PoC VR training scenario was developed to implement and verify key requirements identified in during phase 2, currently missing in the VR training system at BCG. The developed scenario was designed to simulate a hazardous event involving a battery spill and gas leakage, offering users the opportunity to engage with safety-critical actions in a realistic and controlled virtual environment.

This PoC specifically targeted high- and medium-priority requirements from the original requirement specification, particularly those associated with hazard simulation, protective equipment usage, emergency response, and real-time user feedback.

Emphasis was placed on realism, clear procedural guidance, and immersive interaction, in line with requirements such as FR-B1 (hazard simulation), FR-B7 (PPE usage), FR-6 (audio guidance) and UR-4 (learning by doing).

The design of the PoC scenario was guided by the learning theories and HCI principles. While the learning theories did not contribute to a specific requirement, they helped design the PoC scenario flow. Factors such as active participation, engaging in hands-on actions and identifying hazards were all informed by the Experiential learning theory [35].

Usability principles from HCI literature [37] [40] [43] shaped the development of the user requirements specifically, but also the scenario itself. Requirements such as FR-2 (restart option) and FR-5 (help button) aimed to enhance user guidance and error recovery. These were implemented through labeled controls and in-scenario support features, ensuring a smoother and more intuitive training experience for both experienced and inexperienced VR users.

### 5.3.1 Scripted Scenario Flow

The finalized scripted scenario flow was developed and validated during the early development stage. It was reviewed by a BCG trainer to ensure it reflected realistic procedures and aligned with battery safety training practices. The following sequence outlines the scenario flow that supported the development and implementation of the PoC:

**START**

1. User is in a cleanroom and wearing gloves by default.

2. The system instructs the user to pick up the battery, drop it, and then pick it up again.

3. User drops the battery, picks it up, and places it on the table. This action triggers the transition to the inspection step.

4. The system instructs the user to check the battery for leaks or visible faults. A countdown to hazard escalation begins in the background.

5. The battery begins to leak liquid chemicals. After a few seconds, it emits smoke representing the gas leak. The system simulates this using visual and particle effects.

6. A voice instruction informs the user that a gas leak has been detected and instructs them to put on a gas mask. The mask is highlighted and becomes interactable.

7. User picks up the gas mask and puts it on. This action triggers the

next voice instruction.

8. The system instructs the user to evacuate immediately. A yellow guiding circle appears near the exit, and exit signs are activated to make the evacuation path clear.

9. User evacuates by running to the exit. The system ends the scenario with a completion message and a fade-out effect.

**END**

### 5.3.2 Scenario Design Comparison: Current VR System vs. PoC Implementation

Although the current VR system includes a scene involving a smoking battery, it does not fully reflect the vision described by the stakeholders. The existing scenario primarily represents a step within the cell-to-pack phase of battery production, where completed cells are handled, tested and integrated into modules. However, the interface lacks clarity. For instance, the reset/restart button is present but unlabeled, leaving its function unclear. When the battery start smoking the smoke cover the pop-up instructions leaving the user confused and unable to read what to do next. To provide a clearer overview of these limitations and demonstrate how they were addressed in the PoC, Table 5.4 presents a side-by-side comparison of the two scenarios.

| **Feature** | **Existing Scenario** (Current System) | **New PoC Scenario** (Developed) |
|---|---|---|
| Production Phase | Cell-to-Pack | Raw-to-Cell |
| Hazard Type | Thermal runaway with visible smoke caused by battery drop | Gas leak caused by battery drop |
| Visibility of Hazard | Smoke only | Colored smoke representing invisible gas (e.g., HF) |
| Instruction Type | Written instructions | Written + voice instructions |
| User Guidance | No help feature & unlabeled reset button | Help button & labeled reset/restart button |
| PPE Usage | Not required | User must wear a gas mask (respirator) |
| Progress Tracking | A progress menu is shown between each scenario | A progress menu is shown during the PoC scenario |
| Background Sound | None | Background factory sounds were added. |

**Table 5.4:** Comparison between the existing battery hazard scenario in the current VR system and the developed PoC.

### 5.3.3 Scenario Walkthrough

The scenario begins in a clean room setting, where the user is equipped by default with gloves. The hands are visible through a hand-tracked interface to support immersion, see Figure 5.2.

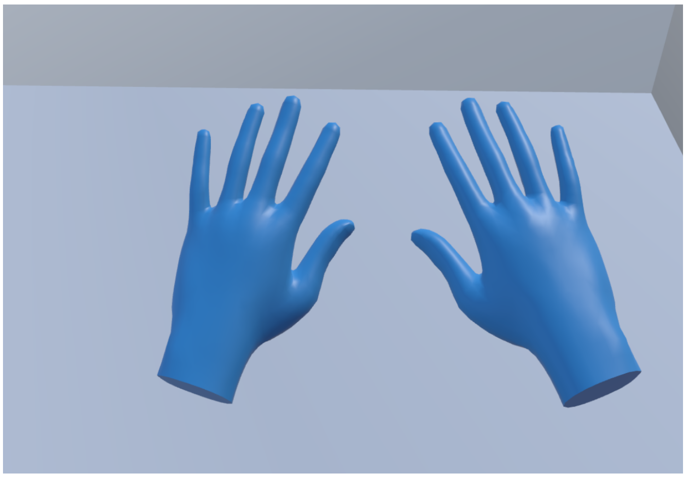


**Figure 5.2:** Initial hand-tracked view showing gloves in the cleanroom environment.

In the clean room, there is a working area including battery cells, PPE and other equipment. The user receives immediate instructions to handle a battery cell, see Figure 5.3, which demonstrates the task environment and on-screen prompts. The battery cell that the user is supposed to interact with is outlined with a yellow color.

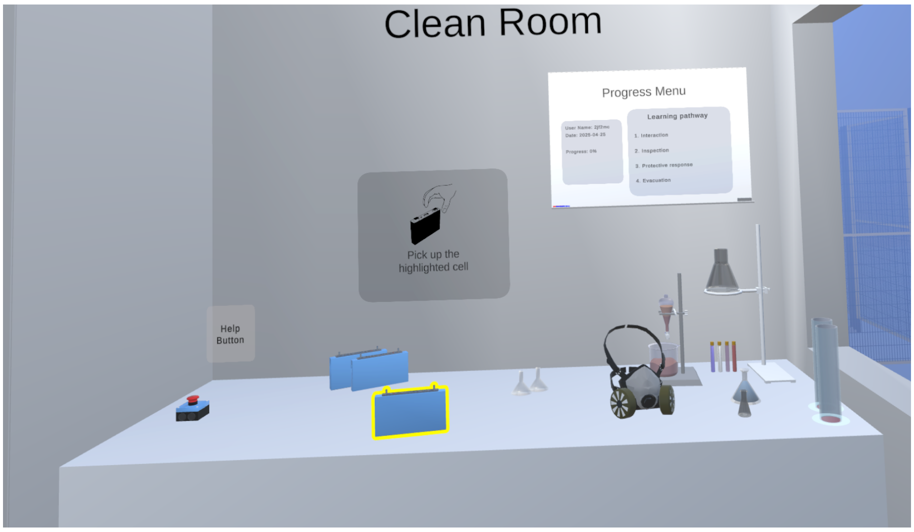



**Figure 5.3:** Cleanroom workspace with an active task prompt to interact with the battery.

As the user proceeds to pick up and drop the battery, as shown in Figure 5.4, the scenario advances through a scripted sequence. Once the battery is dropped, the system triggers a simulated leak and smoke effect to indicate a chemical hazard, see Figure 5.5. These effects were implemented using Unity's particle system. Visual instructions prompt the user to inspect the battery visually, introducing a short moment of inspection before the hazardous situation occurs.

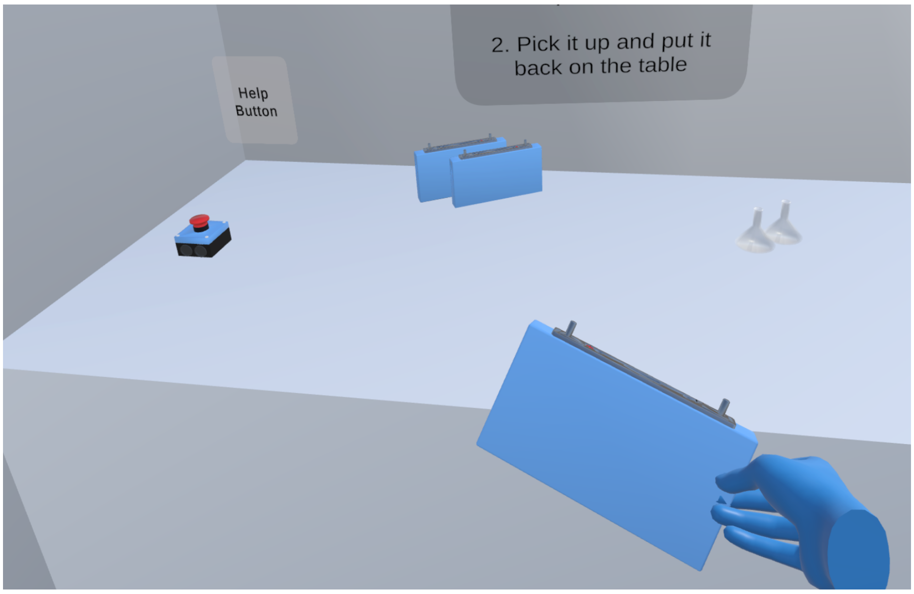



**Figure 5.4:** The user picks up the battery and places it on the table to begin inspection.

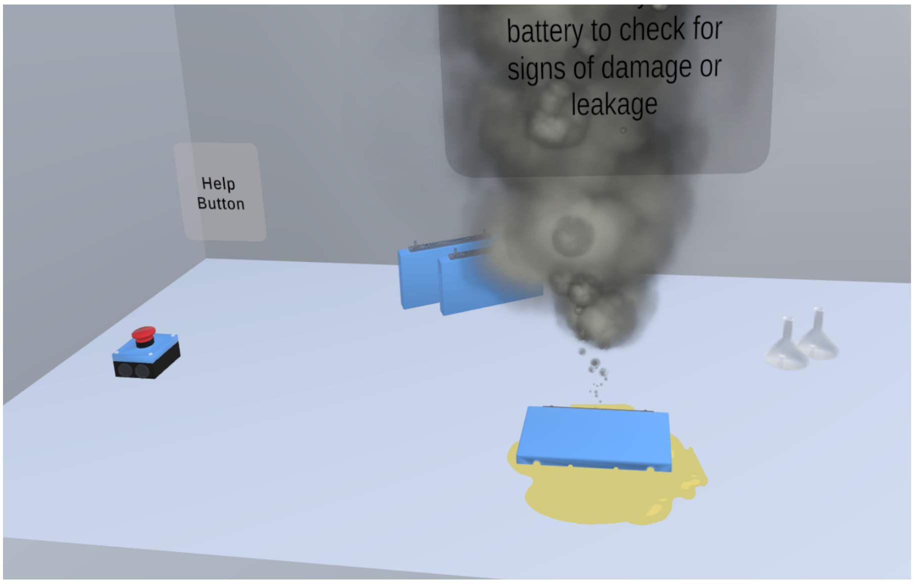
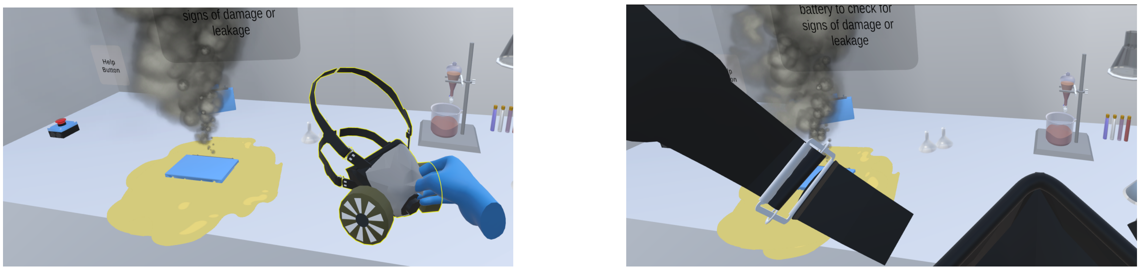



**Figure 5.5:** The battery emits smoke and leaks a chemical to simulate a hazardous event.

As the virtual smoke spreads, voice instructions instruct the user to wear a face mask. The user is guided by visual highlights and object outlines, as seen in Figure 5.6a, where they pick up the mask using their hands and attach it to their face by let go of it. A parallel view during mask interaction is shown in Figure 5.6b, where the smoke is still active in the background.




**(a)** The face mask is highlighted and ready to be equipped.

**(b)** The user is wearing the face mask while the battery cell is still leaking.

**Figure 5.6:** User prepares to wear the mask while the hazardous situation occurs.

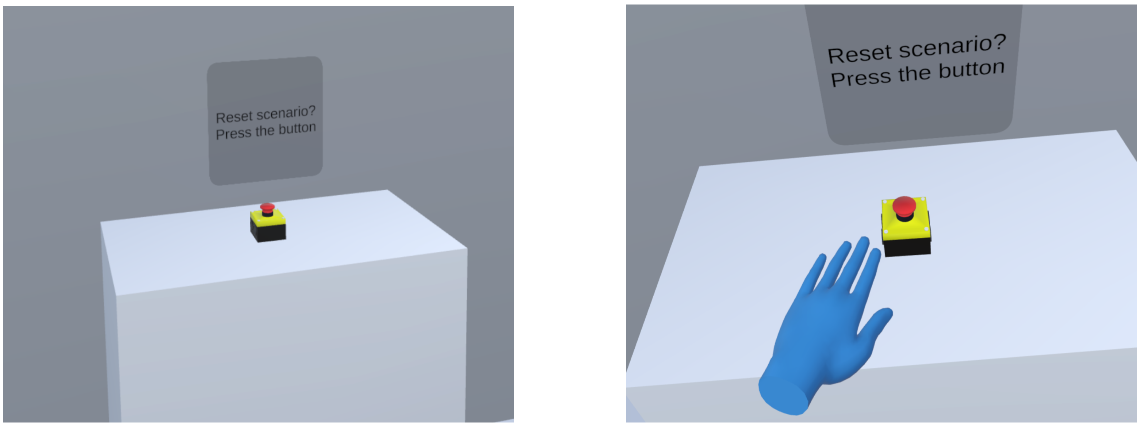
In addition, a help button is available at all stages. Figures 5.7a and 5.7b show how the user can interact with it to seek additional guidance. The user has to press the button and keep their hand on it to view the instruction canvas. When the user stops pressing the button, the instructions disappears and the "Help Button" canvas appears again.


**(a)** User is interacting with the help button.

**(b)** Close-up of help button interaction using hand tracking.

**Figure 5.7:** The help button available throughout the scenario for guidance.

To provide further control, a reset button is also included in the environment, allowing users to restart the scenario from the beginning at any time. This is especially useful in cases of confusion, task failure or discomfort. The reset button is available throughout the scenario, giving users a sense of flexibility without the need for external guidance.

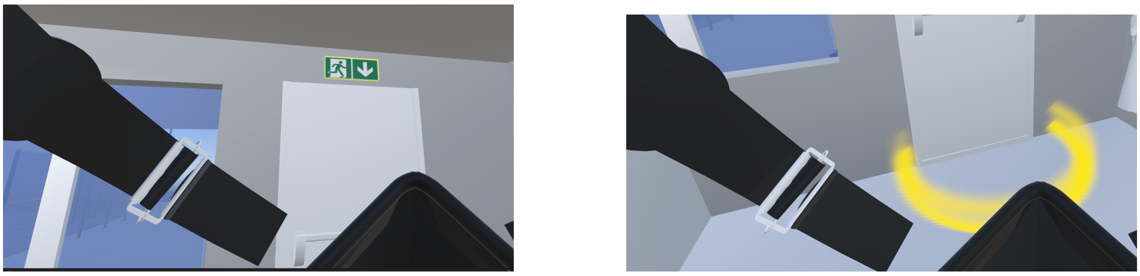



**(a)** Reset button location on the cleanroom table.

**(b)** User interaction with the reset button using hand tracking.

**Figure 5.8:** Reset button functionality allows users to restart the scenario without needing instructor intervention.

After equipping the mask, a final voice cue instructs the user to evacuate. Figures 5.9a and 5.9b show the emergency exit sign being activated and the user approaching the door. This completes the training flow and satisfies evacuation requirements. When the user enters the exit zone, the system presents a white screen fade with the text "Scenario Completed!".



**(a)** The emergency exit sign is activated once the mask is worn.

**(b)** The user proceeds to exit the room, completing the scenario.

**Figure 5.9:** User escapes the hazardous situation with the face mask on.

Finally, Figure 5.10 shows a closer view of the in-scenario progress menu that updates automatically as tasks are completed. The menu also shows the anonymous user name and the current date. The progress menu can also be seen in Figure 5.3, as part of the whole setup.

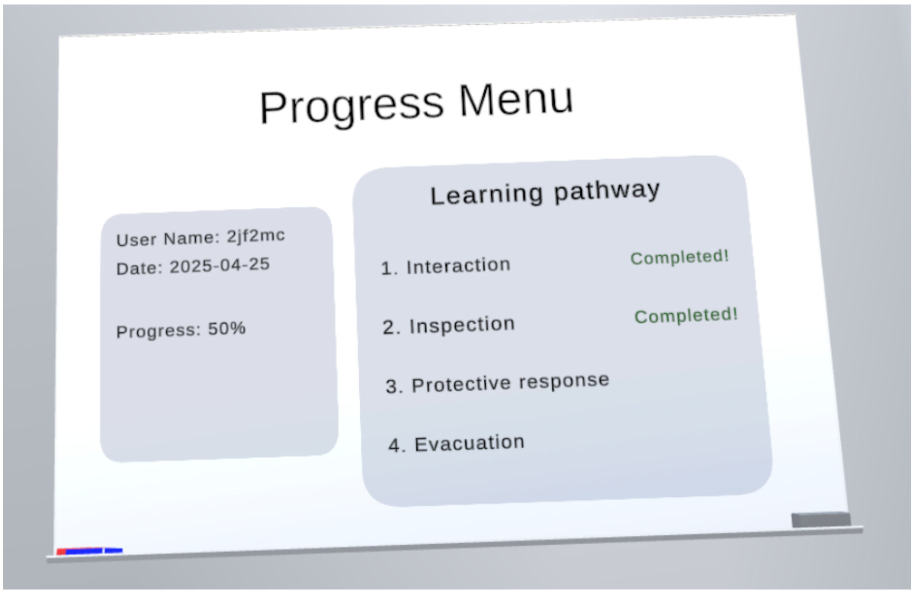



**Figure 5.10:** The dynamic progress menu updates as tasks are completed.

The implementation of the scenario logic relied on two scripts developed specifically for this PoC. The `CellDrop.cs` script governed the core training sequence, handling collision detection, hazard simulation, task progression, visual feedback, and user interaction tracking. The `Head.cs` script was responsible for detecting successful evacuation through the emergency exit zone and concluding the scenario. Full code listings for the developed scripts, including additional implementation details such as the smoke escalation sequence and the interaction setup for the protective face mask, are provided in Appendix A.7.1 and A.7.2 for reference.

### 5.3.4 Requirements Traceability

When the finalized version of PoC was completed and tested, a traceability analysis was performed. To evaluate the alignment between the developed PoC scenario and the original requirement specification. Table 5.5 maps selected high- and medium-priority requirements to their PoC implementations in Phase 3.

Based on the Meta Horizon developer documentation [66] and official specifications [67], the use of the Meta Quest 3 headset contributed directly to fulfilling several system requirements in the PoC scenario due to its technical design and compatibility features. SR-2 was met as the headset supports OpenXR and major XR APIs, making it suitable for cross-platform VR development as required. In addition, SR-3 was fulfilled since the final APK file was well below the 2 GB limit. The headset's adjustable lens spacing (IPD 53–75 mm) and support for glasses-wearing users through an included spacer enabled compliance with SR-8, improving accessibility and user comfort. SR-10 was fulfilled through integration with the Meta XR SDK, which enables full hand tracking.

To provide a comprehensive overview of how each requirement was fulfilled, a requirement matrix is presented in Appendix A.9. It shows which requirements were met through the development of the PoC. It also shows which requirements met by the Meta Quest 3's built-in features, which ones were implemented during the PoC development in Phase 3 and which still need to be tested and verified in Phase 4. Some requirements such as FR-6, UR-1, UR-3, UR-7, UR-11, and NFR-6 need to be tested with users to check how well they work, which can only be confirmed through user feedback and testing. These requirements relate to usability, clarity of instructions, user comfort and training impact. In some cases, some requirements can be confirmed in multiple stages, like FR-7 (casting) and SR-10 (hand tracking), fulfillment involved both hardware support and in-scenario implementation.

Appendix A.8 presents a summary of the requirement coverage by category. Out of 61 total requirements, 37 were addressed either through PoC implementation, hardware capabilities of the Meta Quest 3 or are planned to be verified in Phase 4.

| **Req. ID** | **Category** | **Priority** | **How it was addressed in the PoC** |
|---|---|---|---|
| FR-1 | Functional | High | Visual smoke and audio alerts simulate gas exposure. |
| FR-2 | Functional | Medium | Reset/restart button was added and labeled. |
| FR-5 | Functional | High | Help button added with extra instructions. |
| FR-6 | Functional | High | Two voice instructions were added: gas leak and evacuation. |
| FR-7 | Functional | High | The VR headset supports casting, allowing trainers to observe the session in real time. |
| FR-9 | Functional | Medium | Progress menu updates as user completes each task. |
| FR-12 | Functional | High | Simulated alarm and visual stressors were triggered during hazard escalation to reflect realistic pressure. |
| FR-B1 | Functional | High | Battery drop triggers gas leak simulation using smoke. |
| FR-B2 | Functional | Medium | Leak, smoke and audio simulate escalation. |
| FR-B3 | Functional | Medium | Scenario focuses on raw-to-cell phase, including inspection. |
| FR-B5 | Functional | High | Evacuation triggered after hazard escalation. |
| FR-B7 | Functional | High | User must wear a face mask to proceed. |
| FR-B8 | Functional | High | The progress menu displays "Complete" when a critical step is successfully performed. |

| Req. ID | Category | Priority | How it was addressed in the PoC |
|---|---|---|---|
| UR-1 | User | High | Short, linear scenario with clear guidance. |
| UR-2 | User | High | Help button added for in-scenario support. |
| UR-3 | User | High | Immediate visual/audio cues tied to correct actions. |
| UR-4 | User | Medium | Scenario can be restarted using the reset button. |
| UR-5 | User | Medium | Progress menu updates in real-time. |
| UR-6 | User | High | Followed GTC Trainers description when implementing the scripted scenario flow. |
| UR-7 | User | Medium | Scenario is designed to be short and focused. It took on average 50 seconds to complete the scenario during the final testing. |
| UR-11 | User | High | Cleanroom setting with minimal movement required. |
| NFR-4 | Non-Functional | High | English is the primary language used throughout the scenario. |
| NFR-6 | Non-Functional | Medium | Instructions shown step-by-step without overlap. |
| NFR-7 | Non-Functional | Medium | The scenario starts in under 3 seconds. |
| NFR-9 | Non-Functional | Medium | Scenario can be restarted in less than 5 seconds after failure. |
| NFR-12 | Non-Functional | High | The scenario uses hand tracking and stationary gameplay to avoid motion sickness. |
| SR-1 | System | High | The PoC was developed using Unity with OpenXR support. |
| SR-2 | System | High | The PoC was compatible with Meta Quest 3 during testing and deployment |
| SR-3 | System | High | The apk file is around 100 MB which is less than 2 GB. |
| SR-4 | System | Medium | Room boundaries and safe play space were set up before testing, aligned with headset calibration features. |
| SR-5 | System | High | No personal data is collected or stored in the PoC. |
| SR-6 | System | Medium | The PoC is installable via sideloaded APK on Meta Quest 3. |
| SR-7 | System | High | The system runs on Meta Quest 3 with 12 GB RAM and XR2 chipset. |

| Req. ID | Category | Priority | How it was addressed in the PoC |
|---|---|---|---|
| SR-8 | System | High | Headset supports glasses-wearers with adjustable lenses and glasses spacer. |
| SR-10 | System | High | Meta XR SDK used to enable full hand tracking. |
| SR-11 | System | High | The PoC is designed to have a minimal head and body movement from the user. |
| DR-2 | Data | High | All data will remain anonymous. |

**Table 5.5:** Traceability of selected high- and medium-priority requirements addressed in the PoC scenario. Full descriptions of each requirement ID are available in the requirement specification document at `https://doi.org/10.5281/zenodo.15355792`.

## 5.4 Results from Phase 4

Phase 4 represents the final evaluation stage of the study, where the improved VR training prototype was tested with a new group of participants. This phase focused on assessing the training system's real-world effectiveness and user experience. The results in this section cover task completion performance, survey responses, feature-specific insights and final requirement validation.

### 5.4.1 Measured Real Task Performance

In Phase 4, a controlled user study was conducted to evaluate participants' ability to perform a real-world emergency response task after undergoing either VR-based or text-based training. Participants were divided into two groups: a test group (VR-trained) and a control group (text-trained). The task required participants to handle a simulated battery incident involving PPE, hazard recognition and evacuation.

The performance of participants during the real-world battery safety task in Phase 4 was evaluated across key safety-critical actions and completion time. As shown in Table 5.6, participants in the test group (VR-trained) consistently outperformed those in the control group (text-trained) in all task steps. The VR group demonstrated higher success rates in inspecting the battery, reacting to the alarm, and completing the evacuation correctly. Notably, all VR-trained participants used the gas mask appropriately and evacuated without error. The average task completion time for this group was also significantly shorter (64 seconds) compared to the control group (108 seconds), suggesting increased confidence and familiarity with the task flow.

| Metric | Text-based | VR-trained |
|---|---|---|
| Inspection OK | 57% | 86% |
| Gas Mask OK | 100% | 100% |
| Alarm OK | 86% | 100% |
| Evacuated OK | 86% | 100% |
| Avg. Time (s) | 108 | 64 |

**Table 5.6:** Summary of real-world task performance comparison between text-based and VR-trained groups.

### 5.4.2 Survey Responses

#### Quantitative Results

12 participants were involved in Phase 4, all drawn from a student background. The age distribution skewed younger, with 66.7% of participants under the age of 25. None of the participants reported previous experience with battery management, and only 2 individuals had received any prior training in battery safety. 91.7% of Phase 4 participants reported previous experience using VR, mostly for gaming or entertainment purposes. In contrast, only 47.9% of BAF employees and 53.3% of HVE students in Phase 1 had prior VR exposure. The full demographic details for these participants are listed in Table 4.3.

The results from the Phase 4 survey are presented as a bar chart with error bars representing the SD, as shown in Figure 5.11, for easier comparison with Phase 1. The results indicate generally positive responses to the VR battery safety training. Participants rated their experience across ten thematic areas using a 5-point Likert scale. Half of the themes received mean scores of 4 or above, particularly in areas such as Immersion, Safety Recall, and Educational Value, suggesting that participants found the training engaging, memorable, and informative. Notably, themes such as Confidence and Understanding Risks scored somewhat lower, indicating room for improvement in how the training fosters a sense of preparedness and comprehension of safety-related knowledge.

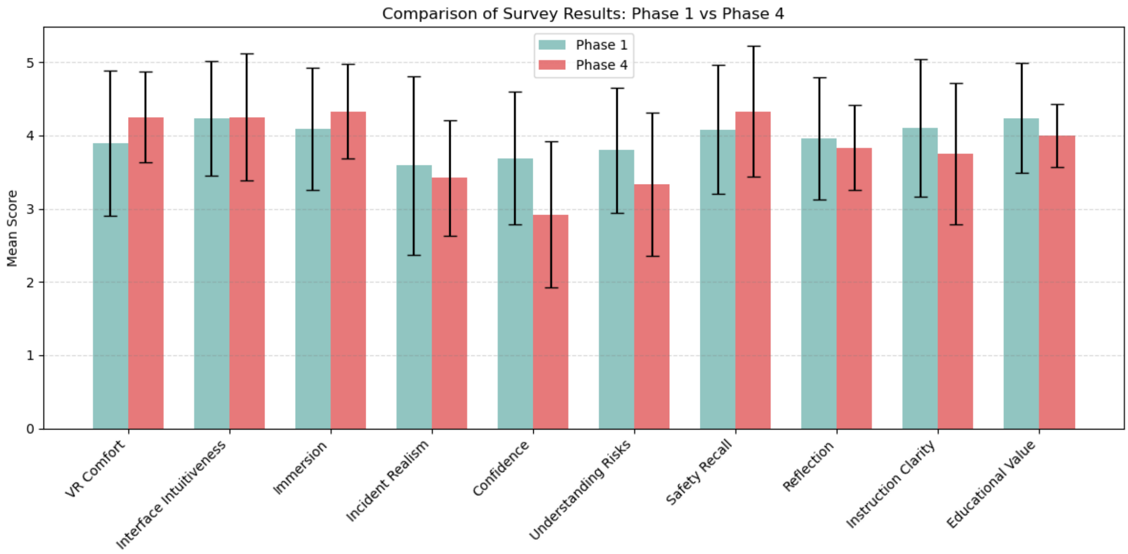



**Figure 5.11:** The mean scores with SD of quantitative survey responses from Phase 1 vs Phase 4, categorized into themes.

A detailed comparison of survey results between Phase 1 and Phase 4 is presented in Table 5.7. The Phase 1 values are average calculations between the two groups. The data show a mixed pattern of improvements and regressions across the ten evaluated themes. Notable increases in mean scores were observed for VR Comfort (+0.36), Immersion (+0.24), and Safety Recall (+0.25), indicating improved user experiences in terms of comfort, engagement, and retention of safety procedures. Instruction Intuitiveness also showed a marginal increase (+0.02), suggesting slight enhancements in interface usability.

In contrast, several learning-related themes experienced declines. Confidence dropped by 0.77 points, the most significant negative shift, while Understanding Risks and Instruction Clarity decreased by 0.47 and 0.35 points respectively. These results may point to gaps in how effectively the revised training communicates complex safety information and builds user confidence. Educational Value also declined slightly (–0.24) but remained high overall.

SD changes add another dimension to the findings. Decreases in SD for themes such as Incident Realism (–0.43), Reflection (–0.26), and Educational Value (–0.32) suggest greater consistency in how participants experienced these aspects of training. Meanwhile, increased SDs in Confidence (+0.09) and Understanding Risks (+0.13) indicate more variability in how different participants perceived these critical learning outcomes. The details of these quantitative survey results are summarized in Appendix A.4.2.

| Theme | Mean (P1) | SD (P1) | Mean (P4) | SD (P4) | Change Mean | Change SD |
|---|---|---|---|---|---|---|
| VR Comfort | 3.89 | 0.99 | 4.25 | 0.62 | +0.36 | −0.37 |
| Interface Intuitiveness | 4.23 | 0.78 | 4.25 | 0.87 | +0.02 | +0.09 |
| Immersion | 4.09 | 0.83 | 4.33 | 0.65 | +0.24 | −0.18 |
| Incident Realism | 3.59 | 1.22 | 3.42 | 0.79 | −0.17 | −0.43 |
| Confidence | 3.69 | 0.91 | 2.92 | 1.00 | −0.77 | +0.09 |
| Understanding Risks | 3.80 | 0.86 | 3.33 | 0.98 | −0.47 | +0.12 |
| Safety Recall | 4.08 | 0.88 | 4.33 | 0.89 | +0.25 | +0.01 |
| Reflection | 3.96 | 0.84 | 3.83 | 0.58 | −0.13 | −0.26 |
| Instruction Clarity | 4.10 | 0.94 | 3.75 | 0.97 | −0.35 | +0.03 |
| Educational Value | 4.24 | 0.75 | 4.00 | 0.43 | −0.24 | −0.32 |

**Table 5.7:** Detailed comparison of Phase 1 (P1) and Phase 4 (P4) survey results. Positive mean changes (green) indicate improved perceptions. Red highlights indicate negative change or increased variability.

To address the possibility of skewed means in the survey results, a comparison between the median scores of Phase 1 and Phase 4 was conducted, see Figure 5.12. Increases were observed in Interface Intuitiveness and Safety Recall, while Confidence and Understanding Risks showed slight decreases, which shows similar directional trends as the mean-based analysis.

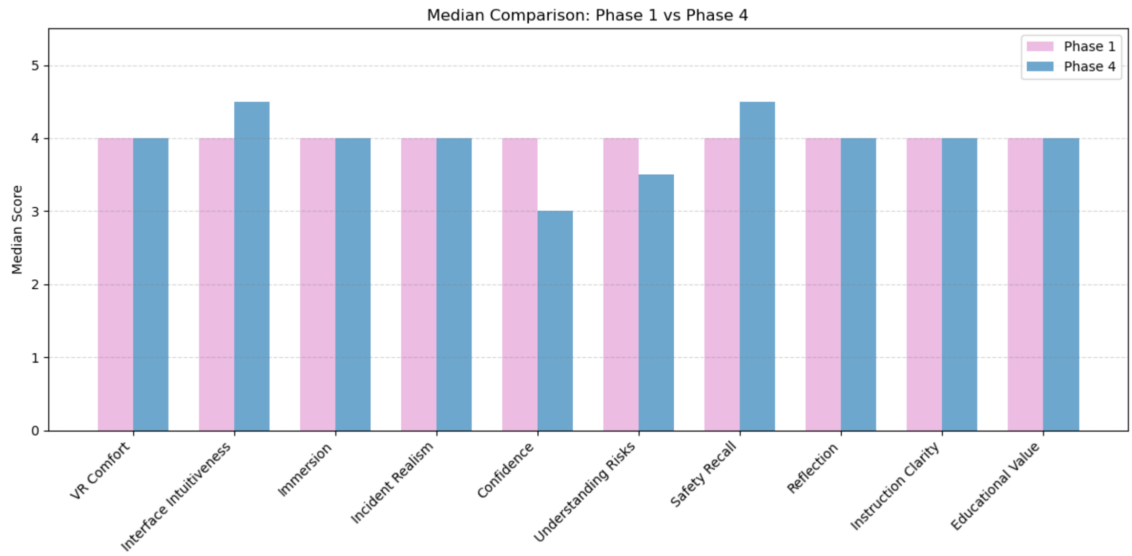



**Figure 5.12:** Median scores of quantitative survey responses from Phase 1 and Phase 4, categorized into themes.

To better understand participants' experience with the VR training system, one specific survey question addressed whether the system provided clear and immediate feedback. In Phase 1, 79.5% of participants responded "Yes", 12.8% said "No" and 7.7% selected "Sometimes." In Phase 4, the positive responses increased slightly to 83.3%, with "No" responses decreasing to 8.3%. "Sometimes" increased slightly to 8.3%. These results indicate that the system was generally successful in delivering timely and understandable feedback across both phases, though minor variations remain.

#### PoC-specific Features

In addition, a set of PoC-specific survey items was included to evaluate the use and perceived value of new features such as the Help and Restart buttons introduced in the updated VR training scenario. Only one participant (8.3%) reported using the Help button, and none used the Restart button during the session. This may indicate that most participants found the scenario sufficiently intuitive or did not feel the need for additional support during the session. The participant who used the Help function rated the helpfulness was 4 out of 5. Notably, all participants (100%) agreed that a Help button would be important in a more complex training system with multiple tasks. The scenario guidance was overall rated as clear, with an average clarity score of 4.42 out of 5.

#### NASA-TLX Survey Results

As previously described in Section 4.1.3, the NASA-TLX is a widely used workload assessment tool that evaluates perceived workload across six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration [58].

Table 5.8 presents the mean scores and SD for each dimension based on participant responses in both Phase 1 and Phase 4. These values reflect how demanding or challenging users perceived the VR training to be across the two phases. Changes in mean and SD between the phases are also included to show differences in user-perceived workload. In Phase 4, the average mental demand score was 40.83, compared to 70 in Phase 1 and physical demand averaged 21.25 in Phase 4, down from 35.83 in Phase 1. Effort also decreased, from 51.67 in Phase 1 to 30.42 in Phase 4. Temporal demand increased slightly from 36.67 in Phase 1 to 40 in Phase 4. Similarly, frustration increased from 25.83 to 29.17. The performance dimension rose from 33.33 to 45. The overall workload score, calculated as a weighted composite across all dimensions, was 47.95 (SD = 15.13) in Phase 1 and 39.25 (SD = 18.43) in Phase 4, indicating a decrease in the general workload perception between the two phases.

| Dimension | Phase 1 | | Phase 4 | | Change | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| **Mental Demand** | 70.00 | 7.07 | 40.83 | 16.63 | −29.17 | +9.56 |
| **Physical Demand** | 35.83 | 18.55 | 21.25 | 11.10 | −14.58 | −7.45 |
| **Temporal Demand** | 36.67 | 24.22 | 40.00 | 22.96 | +3.33 | −1.26 |
| **Performance** | 33.33 | 33.26 | 45.00 | 34.70 | +11.67 | +1.44 |
| **Effort** | 51.67 | 18.86 | 30.42 | 15.29 | −21.25 | −3.57 |
| **Frustration** | 25.83 | 24.58 | 29.17 | 24.76 | +3.34 | +0.18 |
| **Overall Score** | 47.95 | 15.13 | 39.25 | 18.43 | −8.70 | +3.30 |

**Table 5.8:** Average and SD for NASA-TLX scores in Phase 1 and Phase 4, with highlighted changes (green = improved, red = worsened).

### Open-ended Survey Responses

Thematic analysis of the open-ended responses in Phase 4 revealed five recurring themes: scenario length and variety, realism and environment, instruction feedback and timing, technical and behavioral bugs, and positive impressions, see Table 5.9. Several participants requested longer or more varied scenarios to better reflect real-world risks and improve training depth, echoing earlier participant comments made in Phase 1. Suggestions for improved timing and clarity of instructions remained consistent, highlighting a shared need for dynamic instructional feedback. One suggestion was to make the canvas instructions always follow the user's eyes, regardless of where they look, to always have the instructions close-by. Reports of technical or behavioral issues, such as object interaction inconsistencies, were less frequent than in Phase 1. Positive impressions were again common in Phase 4, reinforcing earlier observations that VR training is generally perceived as more engaging and effective than traditional classroom methods.

| Theme | Description |
|---|---|
| Scenario length and variety | Requests for longer or more diverse scenarios to increase training depth and realism. |
| Realism and environment | Feedback related to how closely the VR environment resembled the real world. |
| Instruction feedback & timing | Suggestions to improve audio prompts, response timing, and instructional clarity. |
| Technical and behavioral bugs | Reports of issues with object interactions or system behavior, such as unexpected item drops. |
| Positive impressions | General praise for the training's usefulness, engagement, and clarity compared to traditional methods. |

**Table 5.9:** Themes derived from Phase 4 open-ended survey responses

### 5.4.3 Observations from Phase 4

During the Phase 4, several patterns were observed among participants. Many participants appeared stressed while going through the VR scenario and often struggled to read the written instructions thoroughly. However, despite this, all participants consistently followed the audio instructions. The audio cues seemed to be more effective in capturing attention, since no participant struggled with understand them. Although several participants commented that the PoC scenario felt relatively short, they also expressed that it was impactful and to the point. Furthermore, participants who used the VR training reacted more quickly and confidently during the real measured task. In contrast, those who received only text-based training took more time to think through their actions. While they possessed the theoretical knowledge, they were often unsure how to apply it in practice. On the other hand, those trained in VR appeared more decisive and found the real measured task more straightforward. Participants also appreciated that the task felt realistic and different from a typical experiment, making it more engaging.

### 5.4.4 Final Requirement Validation

This subsection presents the final validation of the requirements based on the results from the user testing conducted in Phase 4. The goal was to evaluate whether the implemented features, behaviors, and user experiences in the PoC actually fulfilled the requirements as intended. While Phase 3 focused on identifying which requirements were addressed through development or hardware capabilities, Phase 4 involved user-centered testing to assess performance, usability and training effectiveness. This was important for validating requirements related to immersion and knowledge transfer, elements that cannot be confirmed through technical implementation alone.

For instance, Phase 4 testing confirmed whether users could complete the training without prior VR experience (UR-1), respond to emergency cues such as gas leak alerts (FR-6), receive and interpret real-time feedback (UR-3) and feel confident applying knowledge in real-world contexts (UR-6). These requirements were assessed through survey responses, behavioral observations and performance metrics.

A requirement was considered fulfilled if it was either (1) implemented and functionally demonstrated in the PoC scenario, (2) supported by the technical capabilities of the Meta Quest 3 headset, as confirmed through official documentation, or (3) validated through user testing and feedback in Phase 4. Requirements that were dependent on user experience, behavior or learning effectiveness were not assumed fulfilled through implementation alone. Table A.8 summarizes this mapping, and Table 5.10 provides a color-coded overview: green indicates fulfilled, yellow indicates not prioritized or partially scoped, and red indicates out of scope.

A majority of the high-priority functional, battery-specific, system and user requirements were successfully fulfilled. These included core safety interactions such as enforcing protective equipment (FR-B7), responding to hazard cues (FR-1, FR-6) and completing evacuation procedures (FR-B5). Requirements such as tracking and logging user interactions (FR-3) and inclusion of a tutorial stage (FR-4) were not prioritized and therefore not implemented or validated. Requirements such as interacting with battery tools were irrelevant for the PoC scenario and therefore marked as out of scope. System-level requirements such as hardware compatibility (SR-2), file size constraints (SR-3), and hand tracking functionality (SR-10) were also validated through the Meta Quest 3's integrated features.

User-centered requirements such as task clarity (UR-1, UR-3), usability (UR-2), and scenario design (UR-6, UR-11) were supported by both survey results and real-world task observations. For instance, high average ratings for comfort and immersion reflected an intuitive user experience. Although the help and restart buttons (FR-2, FR-5) were implemented, they were rarely used—likely due to the scenario's short duration (53 seconds on average), which also fulfilled UR-7 (scenario should be under 20 minutes).

| No. | FR | FR-B | NFR | SR | UR | DR |
|---|---|---|---|---|---|---|
| 1 | green | green | yellow | green | green | yellow |
| 2 | green | green | yellow | green | green | green |
| 3 | yellow | green | yellow | green | green | yellow |
| 4 | yellow | red | green | green | green | yellow |
| 5 | green | green | yellow | green | green | yellow |
| 6 | green | yellow | green | green | green | yellow |
| 7 | green | green | green | green | green | gray |
| 8 | yellow | green | red | green | yellow | gray |
| 9 | green | gray | green | red | yellow | gray |
| 10 | yellow | gray | red | green | yellow | gray |
| 11 | yellow | gray | red | green | green | gray |
| 12 | green | gray | green | gray | gray | gray |
| 13 | yellow | gray | gray | gray | gray | gray |

**Table 5.10:** Color-coded requirement matrix. Green = fulfilled, yellow = not prioritized, red = out of scope, gray = non-existent.

Some requirements marked in yellow and red were considered less critical or outside the scope of this PoC (e.g., multilingual support). These include several non-functional and data-related requirements that would be more relevant in a full-scale training system with several scenarios.

Overall, the implementation achieved a high degree of alignment with the original requirement specification. Most high-priority requirements identified in Phase 2 were fulfilled, either through development, hardware capabilities, or confirmed usability, considering the non-prioritization of the data requirement category. Table 5.11 summarizes the percentages of the fulfilled requirements, excluding the red-coded out of scope requirements.

| **Category** | **Ratio** | **Percentage** |
|---|---|---|
| FR | 7/13 | 54% |
| FR-B | 6/7 | 86% |
| NFR | 5/9 | 56% |
| SR | 10/10 | 100% |
| UR | 8/11 | 73% |
| DR | 1/6 | 17% |

**Table 5.11:** Percentage of requirements fulfilled by category in the PoC implementation.

# 6

# Discussion

This chapter interprets the study's findings in relation to the research questions. It explores how a structured, requirement-driven approach (RQ1) and HCI design principles (RQ2) shaped the evaluation of a VR-based safety training system for the battery manufacturing industry. The chapter also discusses methodological reflections, limitations, and implications for future research and practice.

## 6.1 Addressing RQ1

**RQ1:** *What requirements need to be considered for VR to be effectively used as a training tool in industrial settings, with battery safety training as a use case?*

To make VR an effective training tool in industrial settings, it is important to define what the system must achieve from both a technical and user-centered perspective. Success in this context depends not only on the system's ability to run smoothly and simulate realistic scenarios, but also on its ability to support learning, usability, and safety. In this study, a structured set of requirements was used to guide how VR-based safety training systems were developed and evaluated. Based on that, a PoC scenario was created, which was tested on users to see how well it worked in practice.

The final specification consisted of 61 requirements grouped into six categories: general functional, battery-specific functional, non-functional, system, user, and data requirements. This categorization was based on IEEE 830-1998 [45] and Lauesen's framework [44]. It allowed balance between software engineering best practices and user-centered design. Each category played a different role in shaping the VR training system. Using these six categories provided a structured design approach that balanced technical, experiential and learning-oriented requirements. By separating the requirements into distinct but interrelated categories, it became easier to identify gaps, avoid redundancy and ensure that no critical dimension was overlooked. This helped minimize the risk of building a system that is technically functional but ineffective as a training tool or impractical to implement at scale.

Prioritization was introduced to help manage scope while ensuring that the most critical elements of the VR-based safety training system were implemented first.

Categorizing each requirement as high, medium or low facilitated the focus on features that directly supported core outcomes such as safety, usability, learning effectiveness and system reliability. High-priority requirements were set to 100% fulfillment because they formed the foundation of a safe and pedagogically meaningful experience. Without them, the system would not meet its intended goals. Medium-priority requirements were set at a 70-80% fulfillment threshold. While they enhance the user experience, they are not strictly necessary for the system to function. Low-priority requirements were considered desirable but optional. The percentage thresholds served as a valuable guideline during both development and evaluation. It gave a clear way to measure progress and kept the development focused on what mattered most, without getting stuck trying to build everything at once [56]. This helped make informed trade-offs when balancing time, technical constraints and user needs. It also made it easier to identify which features could be postponed for future iterations without affecting the core goals of safety and learning.

In the PoC, 40 requirements were implemented, with a focus on general and battery-specific functional requirements. These categories were prioritized because they addressed both the core functionality of the VR system and the specific safety risks associated with battery manufacturing. The general functional requirements ensured that the training environment supported basic interactions, feedback and scenario flow. These elements are essential for usability and learning. At the same time, the battery-specific requirements captured critical domain knowledge, such as handling hazardous materials and responding to critical events. User requirements were also prioritized during development to ensure the training experience was accessible, easy to navigate and supported strong knowledge retention, especially for novice VR users. Focusing on these three categories helped create a PoC that was sufficient for evaluation and validation of the requirement specification.

Some requirements were intentionally deprioritized in the PoC phase due to scope and technical constraints. For example, most of the data requirements were not implemented, including DR-3 (exporting logs in CSV/JSON format), DR-4 (logging user interactions) and DR-6 (tracking how many times a user repeats a scenario). These requirements are still highly relevant for future development, but were not necessary to evaluate the PoC features. The PoC covered 100% of all system requirements and 86% of battery-specific ones. Despite not including every possible feature, the PoC still delivered strong user outcomes. This suggests that in early-stage systems, prioritization is more valuable than trying to implement a large number of features. It reinforces the idea that a well-scoped, focused design leads to better results than one that aims for completeness too early. In this way, prioritization became a practical and effective tool for supporting iterative, user-centered development.

When the PoC was tested in Phase 4, the results clearly showed the impact of using a prioritized requirement specification. Participants who trained in VR performed better than the control group in every safety-critical task. They were quicker to put on gas masks, responded more effectively to hazards, and evacuated more efficiently. On average, they completed tasks 40 percent faster and remembered procedures more accurately. These outcomes suggest that taking a requirement-focused approach really made a difference. For instance, the inclusion of audio cues (FR-6) helped users react faster, while clear task flow and support for first-time users (FR-9, UR-1) made the training feel more intuitive. It is clear that VR can offer more than just a new or exciting experience. When the design is built on a strong pedagogical and usability principles, VR has the potential to deliver a training experience that is not only engaging but also highly effective.

The PoC also introduced new features not present in the BVT, such as a help button and a clearer reset button. The help button was only used by 8.3% of participants, although 100% of the participants agreed it would be useful in more complex training scenarios. This aligns with user requirements like UR-2 and FR-5, which focus on providing support, autonomy and recovery from errors. This also indicates that the current scenario was perhaps too simple to require support, but the existence of support mechanisms increased perceived system usability. Beyond usability, the PoC also led to significant improvements in training outcomes, which are discussed in more detail in Section 6.2.

This requirement specification offers a foundation for developing future VR safety training systems. Its structure can be reused or adapted to suit a range of industrial settings. While the framework is designed to be general, the battery production use case shows how it can also be adapted to specific domains. In that context, domain-specific functional requirements were added to address the particular risks and procedures of battery manufacturing. This illustrates that even though the overall structure applies broadly, certain elements need to be customized to fit the demands of the environment. For trainers and stakeholders can use the specification as a tool to identify which features are essential from the beginning and which can be added later. To ensure the requirements align with actual training demands, it's important for developers to engage early with both end-users and safety professionals. Linking goals to specific requirements also made it easier to evaluate the system's impact and provides a strong basis for future improvements.

## 6.2 Addressing RQ2

**RQ2:** *How can HCI principles impact engagement and effectiveness in VR-based training?*

The design of the PoC scenario included several HCI principles to improve user experience and facilitate knowledge transfer. Features such as hand tracking, outlining objects, audio feedback and simplified UI elements were all implemented to lower interaction barriers, reduce user error and promote flow [10]. The results across Phase 1 and 4 suggest that the integration of HCI principles positively influenced user engagement, but also exposed critical areas for further refinement.

Interface intuitiveness was rated highly in both phases (Phase 1 mean = 4.23 and

Phase 4 mean = 4.25), indicating that factors such as hand tracking and object outlining continues to be successful. VR comfort received a higher increase, from 3.89 to 4.25, with a solid drop in standard deviation, suggesting a more comfortable experience overall in Phase 4. This is likely due to the simplicity of the PoC design. It was stationary, brief and had minimal visual clutter, aimed to reduce motion sickness and general discomfort.

Feedback and instruction clarity was a recurring theme in both phases. 83.3% in Phase 4 said feedback was clear, slightly higher than 79.5% in Phase 1, but the instruction clarity decreased. There were still requests and suggestions of clearer instructions. This contradiction may be due to the difference between how user received directions and how the system responded to user actions. As Li [39] notes, intuitive and responsive system feedback can improve perceived usability, even if initial instruction delivery remains a challenge. Since the implemented audio instructions were direct feedback of a user action, it still indicates that the users perceived better feedback quality, but that the instructions before actions can be further improved.

Immersion scores improved in Phase 4, suggesting a more consistently engaging experience. Participants were positive to the audio cues, the consequence simulation (leak, smoke and alarm) and that the scenario flow could be followed through a task progression board. There were less comments about physics or object behavior in Phase 4. Yet, the incident realism average decreased. This could be due to the difference in battery management and safety training experience. Most participants in Phase 1 were, to some extent, familiar with battery handling and possible consequences, while almost no participants in Phase 4 had previous experience with batteries. This may have affected how realistic the incident seemed to the inexperienced participants.

Confidence, understanding risks, reflection and educational value were all themes that decreased in average, despite successful task performance. This gap suggests that the PoC scenario, while straightforward, lacked the complexity or reflection depth needed for the participants to feel confident in handling batteries and reflect about their experience. One possible explanation lies in the intentional simplification of the VR scenario to enhance usability and reduce cognitive load. While the HCI design improved user experience and interaction, it may have unintentionally reduced the perceived seriousness of the simulated risks. This highlights the need for a training system that includes multiple scenarios addressing different aspects of battery handling and safety, to build confidence and deep understanding.

Despite lower perceived confidence and risk understanding, safety recall improved to an average of 4.33. This suggests that users may not feel confident, but their physical memory of actions were still retained. This outcome is aligned with existing research on procedural memory and embodied cognition in VR, which suggests that repeated physical interaction, even without deep reflection, can lead to effective encoding of action-based skills [40].

The NASA-TLX survey was used to evaluate perceived workload across six dimensions. The results suggest that the PoC scenario reduced user workload overall, but also introduced new trade-offs that have implications for learning effectiveness and scenario design. Mental demand, physical demand and effort all decreased in Phase 4, indicating that users found the system easier to understand, interact with and physically tolerate. From a usability perspective, these changes reflect successful application of HCI principles, such as smooth interaction, guided task flow and clean and simple interface. These are all known to reduce user workload and enhance task performance in immersive systems [38]. Since the results could be interpreted as low-to-moderate, based on a scale from 0-100, there is an indication of cognitive-affective balance, where users focused on the training content rather than struggling with the interface.

Temporal demand and frustration increased slightly in Phase 4, indicating that users felt more time pressure. This may be due to the emergency nature of the scenario, including alarms, gas leak and an evacuation sequence, which naturally introduce urgency. Frustration also increased slightly in Phase 4. This could be linked to the perceived lack of time, or to the short duration of the scenario itself, which may have felt abrupt or incomplete. The performance score increased from 33.33 to 45, which in NASA-TLX terms implies perceived worse performance. Given that all participants performed the scenario with success, this increase may be due to a user misinterpretation of the question. The performance scale in the NASA-TLX survey is reversed, meaning that higher scores correspond to worse outcome. If participants misread or overlooked this reversal, they may have accidentally rated their performance lower than they actually perceived it to be, assuming higher equals better. The result were therefore interpreted with caution.

The individual NASA-TLX dimension scores and the overall workload rating fall on the lower end of the scale, while still not dropping below 20. In the context of VR-based training, the goal would not be to reduce cognitive or temporal demand entirely, but to keep it within a range that supports learning without inducing overload. From an HCI perspective, these results support the importance of balancing challenge and support, ensuring that the system is neither overwhelming nor so easy that it becomes disengaging [38].

The measured real-world task evaluation in Phase 4 showed that all participants in the VR-trained group completed the critical safety tasks correctly, including recognizing the hazard, wearing protective equipment and evacuating the area. The task success rate of inspecting the battery was also higher in the VR-trained group. This alignment between training and real-world actions demonstrates that even a short, focused VR scenario can transfer procedural knowledge effectively. One important note is that all participants used the gas mask, but when it came to reacting and evacuating, the VR-trained group was much faster. As peviously mentioned, the VR-trained groups average time to complete the task was almost twice as fast as the text-based group. This may indicate that hands-on interaction in VR helped prepare participants for independent decision-making, supporting the HCI design goals such as intuitive affordances and experiential feedback, which are known to enhance user task flow and presence in VR environments [40]. Although users reported lower subjective confidence in the survey, their actual behavior suggests that learning outcomes were strong.

## 6.3 Reflections on Method and Design Choices

This study was designed to iteratively evaluate and improve a VR-based training system for the battery industry, through combining HCI principles, user feedback and software engineering methodologies. A key strength of the approach was the use of mixed methods across the data collection phases. By combining quantitative surveys, open-ended responses, group interviews and a real-world performance task, the study was able to provide a more comprehensive view of user experience and training effectiveness than any single method alone would have allowed. This increased the reliability of the findings and the data could be interpreted in greater depth.

The PoC scenario introduced several targeted improvements over the existing VR system, as presented in Table 5.4. The addition of audio instructions alongside written prompts appeared to improve perceived feedback clarity, as previously discussed. This suggests that adding audio guidance helped reinforce user awareness during tasks, particularly under time pressure. However, survey responses also indicated that instruction clarity was still an area needing improvement. Again, this implies that while audio cues were helpful, the initial task framing or timing of instructions could still be improved. In the real-world task, all VR-trained participants used the gas mask appropriately and completed evacuation successfully, indicating successful procedural transfer, which confirms the implementation of PPE interaction. Adding color to the gas leak may have contributed to faster reaction times and improved hazard recognition, as observed in the real-world task performance. Finally, the addition of a factory background sound helped address earlier feedback about a lack of realism. While this specific change was not measured quantitatively, no negative feedback was reported and it may or may not have contributed to the improved immersion score.

An important design choice was the focus on creating a simplified PoC scenario that could be developed and tested within project scope. The choice to make the training stationary, short, and visually simple was guided by both practical limitations and HCI principles related to reducing cognitive load and ensuring accessibility for novice VR users. However, the high proportion of VR-experienced participants in Phase 4 (91.7%) may also have contributed to the improved ratings for VR comfort and interface intuitiveness. Unlike in Phase 1, where nearly half the participants were new to VR, the Phase 4 group likely required less cognitive effort to adapt to the interaction model. This may partially explain the lower mental demand and effort scores in the NASA-TLX survey, as well as the positive responses to immersion and interface intuitiveness. While the scenario design helped improve comfort, immersion and usability, as both survey results showed, it also limited the depth and complexity of the training. This helps explain why participants showed strong behavioral outcomes but lower self-reported confidence and understanding of risk. In short, there was a trade-off between comfort and usability on one hand, and training depth and reflective insight on the other.

The instructional design was also partly guided by several learning theories, introduced in Section 2.4. Kolb's experiential learning model was reflected in how users engaged directly with the hazard and experienced consequences, and had to take action themselves, providing a concrete experience. The SECI model's concepts of internalizing knowledge through action were implemented to some extent, especially in how users physically practiced safety behaviors. Although the duration of the PoC limited the full implementation of these learning theories, they still help explain why procedural retention was strong despite lower self-reported confidence.

Finally, the inclusion of a real-world task in Phase 4 added insights into how the training translated to behavior. The strong performance by VR-trained users validated the effectiveness of the training design. However, participants in the text-based group had to read and remember a broader set of instructions, many of which were not directly relevant to the immediate task. This may have made it harder for them to identify which actions were most critical, highlighting a limitation of traditional learning formats. The task was still limited in scope, since it was just as short as the PoC, and conducted in a controlled environment. Long-term retention, stress handling and on-the-job performance were not measured and remain areas for future work.

### 6.3.1 Threats to Validity

As with many empirical software engineering studies, certain trade-offs were necessary in the design of this research. Stol and Fitzgerald [52] highlight that balancing internal validity, realism, and generalizability is an ongoing challenge in applied research. In this study, the focus was on evaluating a functional PoC in a realistic training context, rather than conducting a fully controlled laboratory experiment.

When interpreting the results from an internal validity perspective, the comparison between Phase 1 and Phase 4 was not a fully controlled experiment. The testing structure was the same, but they differed in participants demographics. The changes in the quantitative survey data may be influenced by these group differences rather than the intervention alone. For instance, many participants in Phase 1 had prior battery experience, while Phase 4 consisted mostly of participants with limited battery knowledge. There was also a difference in prior VR experience between the two groups. In Phase 4, 91.7% of participants reported previous VR experience, while Phase 1 included a higher proportion of first-time users. As a result, some observed improvements may reflect user familiarity rather than improvements in the system itself. Additionally, motivational and novelty effects may have influenced outcomes. Participants in the VR group may have been more engaged due to the inherent novelty and immersive nature of the technology, possibly increasing engagement or performance metrics compared to those in the text-based control group. While such effects are difficult to isolate, future studies should consider more neutral control conditions, for instance video-based instructions, to reduce potential biases in enthusiasm or engagement levels.

Another confounding factor relates to the difference in instructional format between groups. The text-based group read approximately eight pages of diverse battery safety procedures, many of which were not directly relevant to the specific task that was later tested. In contrast, the VR scenario presented a focused, one-minute experience designed explicitly to simulate that task. This may have disadvantaged the control group by introducing higher cognitive load and less targeted preparation. Similarly, the short duration and narrow scope of the PoC scenario may help explain why VR-trained participants performed well but still reported lower confidence and risk understanding. Although effective at building procedural memory, the scenario may not have provided enough reflection to support deeper learning.

In terms of construct validity, there is a potential risk that certain survey questions, particularly in the NASA-TLX survey, were misinterpreted. As previously mentioned, the performance dimension is reversed (higher scores indicate worse performance) and participants may have have mistakenly rated it in the wrong direction. This issue could have skewed the mean performance score in both Phase 1 and Phase 4.

External validity is limited by the scope of the PoC and the size of the real-world task group. While the VR scenario was designed using realistic procedures and equipment, it represented a small part of the broader training needs in battery manufacturing. Furthermore, the relatively small group size and focus on short-term effects limit the generalizability of the outcomes. Replicating the experiment with larger samples and more diverse settings would strengthen the conclusions.

In regards to conclusion validity, while several positive trends were observed, such as improved safety recall and faster performance in the real-world task, not all of these changes were statistically tested due to the limited sample size. Additionally, since no formal pre- and post-test of knowledge was conducted, it was not possible to precisely measure knowledge gain over time. Instead, learning effectiveness was interpreted from user performance and survey responses. Furthermore, the text-based condition, while effective for isolating VR's experiential benefits, may not fully reflect real-world training alternatives. Traditional methods often include interactive classroom sessions or hands-on practice which differ from reading-based formats. Despite these limitations, the use of multiple data sources, including surveys, interviews and observed task behavior, suggests that the conclusions are robust within the scope of the study.

## 6.4 Broader Implications and Future Work

While this research focused on battery manufacturing as a use case, the findings have broader implications for other sectors where risk awareness, procedural training and situational decision-making are critical. Industries such as chemical processing, construction, logistics and healthcare could benefit from adopting similar VR frameworks, especially when training must occur in safe, repeatable and cost-effective environments. The use of VR in these contexts can offer safe and immersive environments that promotes trainees to engage with hazardous procedures and decision-making in ways traditional methods cannot replicate.

From a pedagogical perspective, the results suggest that even short, focused VR scenarios can effectively support procedural memory, engagement and task recall. This suggests that immersive technologies can serve as a complement to traditional classroom or on-site instruction, particularly for onboarding and early-stage skill acquisition. However, the results also highlighted the importance of instructional depth, reflection opportunities and scenario variety to support confidence, conceptual understanding and higher-order learning.

For developers and designers, this thesis provides a reusable template for developing effective VR training systems by presenting a validated requirement specification based on both empirical research and stakeholder input. By mapping user needs and pedagogical goals into categorized and traceable requirements, teams can build training solutions that are more aligned with both industry and learner expectations. In addition, the integration of HCI principles such as intuitive interaction, feedback clarity and cognitive load management proved essential for engagement and usability. For trainers, the study shows how immersive training environments can be used not only for knowledge transfer but also for reinforcing safe behaviors through experiential learning. The prioritization framework also serves as a decision-making tool, enabling trainers to prioritize resources toward the most impactful system features.

This study also contributes to theory in two main areas. First, the project adds to HCI research by applying user-centered design methods in a domain-specific VR context, validating prior findings on the value of simplicity, immersion and interaction clarity. Second, the structured application of requirement engineering principles in an emerging technology domain contributes to the field of software engineering, especially in the design of safety-critical or instructional systems.

There are several opportunities for future work in regards to this project. One future work opportunity is to expand the PoC into a modular training platform with multiple safety scenarios. This would allow for more comprehensive learning outcomes and better reflect the complexity of real-world tasks. Future studies should also aim to include larger participant samples and long-term follow-ups to measure retention, workplace behavior and performance over time. For instance, including knowledge assessments before and after training would strengthen conclusions about learning effectiveness.

Finally, technical improvements such as enhanced visual realism and environmental detail, multiplayer support and personalized feedback mechanisms could further improve the effectiveness and scalability of VR-based training systems. Exploring the possibility of integration with existing learning management systems in industrial settings would also increase real-world applicability and adoption potential.

# 7

# Conclusion

This thesis explored how VR can be effectively used to support safety training in industrial settings, with battery manufacturing as a use case, through a four-phase study. The first phase included a pre-study evaluation of the BVT system involving 78 participants through surveys and group interviews. The second phase included the development of a structured and prioritized requirements specification, grounded in user feedback, stakeholder input and literature. The third phase involved the design and implementation of a PoC scenario based on the specification and in the fourth phase, a controlled experiment with a new set of participants was conducted, comparing the VR-trained group to a text-trained control group using a real-world task assessment. Through the development, testing and evaluation of the BVT system at BCG, the study aimed to understand both the requirements needed for such a training tool and how HCI principles affect engagement and learning outcomes.

In response to RQ1, the study developed a structured requirement specification based on user-centered and software engineering methodologies. These requirements included functional, non-functional, system, data and user-specific needs and were derived from literature search, stakeholder interviews, user observations and participant feedback. This structure ensured that the resulting VR training solution was aligned not only with technical feasibility but also with cognitive and pedagogical training goals, such as safety awareness and procedural memory.

With reference to RQ2, the findings demonstrate that well-applied HCI principles enhance the training experience. By incorporating user-centered design features, such as immediate feedback loops and careful cognitive load management, the VR training scenario kept participants engaged and comfortable, which in turn improved their learning outcomes. In addition, all participants in the VR-trained group completed the critical safety tasks in a real-world evaluation and did so significantly faster than the control group. These results confirm that HCI principles directly contribute to both increased user engagement and the effectiveness of learning in VR. Ultimately, when VR training environment are designed with the user in mind, they become not only more usable, but also more pedagogically powerful, leading to greater knowledge retention and skill transfer in a short period.

However, the study revealed a gap between task performance and perceived confidence and understanding. While participants in the VR group performed well, they did not consistently report feeling confident or deeply reflective about their experience. This highlights a limitation of the PoC's simplicity and the importance of balancing usability with deep and varied tasks where the user can reflect over the scenarios, especially in high-risk environments where reflection and conceptual understanding are critical.

It should be noted that these conclusions are drawn from a specific implementation and sample and should therefore be interpreted within the context of this study's scope. The PoC was tested on a limited group of participants and within a focused battery safety scenario. As such, while the positive trends in usability and immediate learning outcomes are evident, they apply to the conditions defined by the delimitations. Future work can build on this research by exploring longer-term effects and testing the requirement framework in diverse settings.

Overall, this thesis concludes that VR can serve as a highly effective platform for industrial safety training when it is developed according to a well-defined requirement framework and informed by HCI principles. The research contributes a validated set of design requirements for VR training systems and evidence of how these requirements and design choices improve safety training outcomes. By combining requirements engineering with human-centered design, the study offers a practical blueprint for creating engaging and effective VR training experiences. Ultimately, this study shows that when VR training is grounded in user needs and supported by structured design principles, it can be a powerful and effective tool for improving safety in the battery industry and other high-risk sectors.

# Bibliography


[1] Björn Johansson et al. "Challenges and opportunities to advance manufacturing research for sustainable battery life cycles". In: *Frontiers in Manufacturing Technology* 4 (2024). DOI: 10.3389/fmtec.2024.1360076. URL: https://doi.org/10.3389/fmtec.2024.1360076.

[2] Andrea de Giorgio et al. "Adopting extended reality? A systematic review of manufacturing training and teaching applications". In: *Journal of Manufacturing Systems* 71 (2023), pp. 645–663. ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2023.10.016. URL: https://doi.org/10.1016/j.jmsy.2023.10.016.

[3] Invest in Gothenburg. *New Education Centre to Boost Skills Supply to Fast-Growing Battery Industry.* https://www.investingothenburg.com/news/all-news/new-education-centre-boost-skills-supply-fast-growing-battery-industry. Published by Business Region Gothenburg. Accessed: 2025-04-07. Jan. 2025.

[4] Henrik Söderlund et al. "The creation of a Multi-User virtual training environment for operator training in VR". In: *Advances in Transdisciplinary Engineering.* 2024. DOI: 10.3233/atde240163. URL: https://doi.org/10.3233/atde240163.

[5] Paweł Strojny and Natalia Dużmańska-Misiarczyk. "Measuring the effectiveness of virtual training: A systematic review". In: *Computers & Education: X Reality* 2 (2023), p. 100006. DOI: 10.1016/j.cexr.2022.100006.

[6] Oscar Escallada et al. "Assessing Human Factors in Virtual Reality Environments for Industry 5.0: A Comprehensive Review of Factors, Metrics, Techniques, and Future Opportunities". In: *Information* 16.35 (2025), pp. 1–34. DOI: 10.3390/info16010035. URL: https://doi.org/10.3390/info16010035.

[7] D. Scorgie et al. "Virtual reality for safety training: A systematic literature review and meta-analysis". In: *Safety Science* 171 (2024), p. 106372. DOI: 10.1016/j.ssci.2023.106372.

[8] Yuqing Chen et al. "A review of lithium-ion battery safety concerns: The issues, strategies, and testing standards". In: *Journal of Energy Chemistry* 59 (2021), pp. 83–99. DOI: 10.1016/j.jechem.2020.10.017.

[9] Francesco Osti et al. "A VR training system for learning and skills development for construction workers". In: *Virtual Reality* 25 (2021), pp. 523–538. DOI: 10.1007/s10055-020-00470-6.

[10] David Checa and Andres Bustillo. "A review of immersive virtual reality serious games to enhance learning and training". In: *Multimedia Tools and Ap-*

*plications* 79 (9-10 Mar. 2020), pp. 5501–5527. ISSN: 15737721. DOI: 10.1007/s11042-019-08348-9.

[11] Sarune Baceviciute et al. “Investigating the Value of Immersive Virtual Reality Tools for Organizational Training: An Applied International Study in the Biotech Industry”. In: *Journal of Computer Assisted Learning* 38.2 (2022), pp. 470–487. DOI: 10.1111/jcal.12630. URL: https://doi.org/10.1111/jcal.12630.

[12] Julian Abich IV et al. “A Review of the Evidence for Training Effectiveness with Virtual Reality Technology”. In: *Virtual Reality* 25 (2021), pp. 919–933. DOI: 10.1007/s10055-020-00498-8. URL: https://doi.org/10.1007/s10055-020-00498-8.

[13] Guido Makransky and Gustav B. Petersen. “The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality”. In: *Educational Psychology Review* 33 (2021), pp. 937–958. DOI: 10.1007/s10648-020-09586-2.

[14] Raveekiat Singhaphandu and Warut Pannakkong. “A Review on Enabling Technologies of Industrial Virtual Training Systems”. In: *International Journal of Knowledge and Systems Science (IJKSS)* 15.1 (2024), pp. 1–33. DOI: 10.4018/IJKSS.352515. URL: https://doi.org/10.4018/IJKSS.352515.

[15] Heather C. Lum et al. “Virtual Reality: History, Applications, and Challenges for Human Factors Research”. In: *Proceedings of the 2020 HFES 64th International Annual Meeting* 64.1 (2020). Copyright 2020 by Human Factors and Ergonomics Society. All rights reserved., pp. 1263–1270. DOI: 10.1177/1071181320641300.

[16] Paul Milgram and Fumio Kishino. “A Taxonomy of Mixed Reality Visual Displays”. In: *IEICE Transactions on Information and Systems* E77-D.12 (Dec. 1994), pp. 1321–1329. URL: https://www.researchgate.net/publication/231514051_A_Taxonomy_of_Mixed_Reality_Visual_Displays.

[17] Michael A. Gigante. “Virtual Reality: Definitions, History and Applications”. In: *VIRTUAL REALITY SYSTEMS* (1993). All rights of reproduction in any form reserved, pp. 1–14.

[18] Tomasz Mazuryk and Michael Gervautz. “Virtual Reality - History, Applications, Technology and Future”. In: *Institute of Computer Graphics, Vienna University of Technology* (Dec. 1999). URL: https://www.researchgate.net/publication/2617390_Virtual_Reality_-_History_Applications_Technology_and_Future.

[19] Morton L. Heilig. *Sensorama Simulator*. US Patent 3,050,870. Filed: 1961-01-10. Granted: 1962-08-28. 1962. URL: https://patents.google.com/patent/US3050870A.

[20] Scott Fisher et al. “Virtual Interface Environment Workstations”. In: *Proceedings of the Human Factors Society Annual Meeting*. NASA Ames Research Center. Moffett Field, California: Human Factors Society, 1988, pp. 2988–2992. DOI: 10.1177/154193128803200219. URL: https://www.researchgate.net/publication/4709345_Virtual_Interface_Environment_Workstations.

[21] Wade Sisler. *Virtual Environment Reality workstation technology (helmet & gloves)*. https://images.nasa.gov/details/ARC-1992-AC89-0437-6.

NASA ID: ARC-1992-AC89-0437-6. Public domain image from NASA Image and Video Library. Accessed: 2025-03-12. 1992.

[22] Jianghao Xiong et al. "Augmented Reality and Virtual Reality Displays: Emerging Technologies and Future Perspectives". In: *Light: Science & Applications* 10.216 (2021). DOI: 10.1038/s41377-021-00658-8. URL: https://doi.org/10.1038/s41377-021-00658-8.

[23] Divine Maloney, Guo Zhang Freeman, and Andrew Robb. "Social Virtual Reality: Ethical Considerations and Future Directions for an Emerging Research Space". In: *IEEE Conference on Virtual Reality and 3D User Interfaces.* 2021. DOI: 10.1109/VRW52623.2021.00056.

[24] Alexander Skulmowski. "Ethical issues of educational virtual reality". In: *Computers & Education: X Reality* 2 (2023), p. 100023. DOI: 10.1016/j.cexr.2023.100023.

[25] Unity Technologies. *Unity - Real-time development platform.* Accessed: 2024-02-20. 2024. URL: https://unity.com/.

[26] Afzal Hussain et al. "Unity Game Development Engine: A Technical Survey". In: *University of Sindh Journal of Information and Communication Technology (USJICT)* 4.2 (2020), pp. 73–81. ISSN: 2523-1235 (Online), 2521-5582 (Print). URL: https://www.researchgate.net/publication/348917348_Unity_Game_Development_Engine_A_Technical_Survey.

[27] Unity Technologies. *Unity Asset Store - The Best Assets for Game Development.* Accessed: 2024-02-20. 2024. URL: https://assetstore.unity.com/.

[28] Mukkamala S.N.V. Jitendra et al. "A Study on Game Development Using Unity Engine". In: *AIP Conference Proceedings* 2375 (2021), p. 040001. DOI: 10.1063/5.0066303. URL: https://doi.org/10.1063/5.0066303.

[29] Alina-Mădălina Gheorghe, Ileana Daniela Gheorghe, and Ioana Laura Iatan. "Agile Software Development". In: *Informatica Economică* 24.2 (2020), pp. 90–100. DOI: 10.24818/issn14531305/24.2.2020.08. URL: https://doi.org/10.24818/issn14531305/24.2.2020.08.

[30] Samar Al-Saqqa, Samer Sawalha, and Hiba AbdelNabi. "Agile Software Development: Methodologies and Trends". In: *International Journal of Interactive Mobile Technologies (iJIM)* 14.11 (2020), pp. 246–270. DOI: 10.3991/ijim.v14i11.13269. URL: https://doi.org/10.3991/ijim.v14i11.13269.

[31] Henry Edison, Xiaofeng Wang, and Kieran Conboy. "Comparing Methods for Large-Scale Agile Software Development: A Systematic Literature Review". In: *IEEE Transactions on Software Engineering* 48.8 (2022), pp. 2709–2736. DOI: 10.1109/TSE.2021.3069039. URL: https://doi.org/10.1109/TSE.2021.3069039.

[32] Soobia Saeed et al. "Analysis of Software Development Methodologies". In: *International Journal of Computing and Digital Systems* 8.5 (2019), pp. 445–460. DOI: 10.12785/ijcds/080502. URL: http://dx.doi.org/10.12785/ijcds/080502.

[33] Barbara Ostrowski Martin, Klodiana Kolomitro, and Tony C. M. Lam. "Training methods: A review and analysis". In: *Human Resource Development Review* 13.1 (2014), pp. 11–35. DOI: 10.1177/1534484313497947.

[34] Shu-Lun Mak et al. “A Review on Development and Application of Virtual Reality (VR) Training Platform for Testing, Inspection and Certification Industry”. In: *International Journal of Information and Education Technology* 10.12 (2020), pp. 926–931. DOI: 10.18178/ijiet.2020.10.12.1480. URL: https://doi.org/10.18178/ijiet.2020.10.12.1480.

[35] David A. Kolb and Alice Y. Kolb. *The Kolb Learning Style Inventory 4.0: Guide to Theory, Psychometrics, Research & Applications*. Experience Based Learning Systems, Inc., 2013. URL: https://www.researchgate.net/publication/303446688_The_Kolb_Learning_Style_Inventory_40_Guide_to_Theory_Psychometrics_Research_Applications.

[36] Ikujiro Nonaka. “The Knowledge-Creating Company”. In: *Harvard Business Review* 69.6 (1991), pp. 96–104. URL: https://hbr.org/1991/11/the-knowledge-creating-company.

[37] Helen Sharp, Jenny Preece, and Yvonne Rogers. *Interaction Design: Beyond Human-Computer Interaction*. 5th. John Wiley & Sons, Incorporated, 2019.

[38] Joe Cecil et al. “Exploring Human-Computer Interaction (HCI) Criteria in the Design and Assessment of Next Generation VR-based Education and Training Environments”. In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2021. DOI: 10.1109/VRW52623.2021.00144. URL: https://doi.org/10.1109/VRW52623.2021.00144.

[39] Tao Li. “Human-computer interaction in virtual reality environments for educational and business purposes”. In: *Management of Development of Complex Systems* 57 (2024), pp. 112–117. DOI: 10.32347/2412-9933.2024.57.112-117. URL: https://doi.org/10.32347/2412-9933.2024.57.112-117.

[40] Alistair G. Sutcliffe et al. “Reflecting on the Design Process for Virtual Reality Applications”. In: *International Journal of Human–Computer Interaction* 35.2 (2019), pp. 168–179. DOI: 10.1080/10447318.2018.1443898. URL: https://doi.org/10.1080/10447318.2018.1443898.

[41] Sam Van Damme et al. “Impact of Latency on QoE, Performance, and Collaboration in Interactive Multi-User Virtual Reality”. In: *Applied Sciences* 14.6 (2024), p. 2290. DOI: 10.3390/app14062290. URL: https://doi.org/10.3390/app14062290.

[42] Nigel Bevan et al. “New ISO Standards for Usability, Usability Reports and Usability Measures”. In: *Human-Computer Interaction. Theory, Design, Development and Practice*. Ed. by Masaaki Kurosu. Vol. 9731. Lecture Notes in Computer Science. Springer, Cham, 2016, pp. 268–278. DOI: 10.1007/978-3-319-39510-4_25. URL: https://doi.org/10.1007/978-3-319-39510-4_25.

[43] Jozsef Katona. “A Review of Human–Computer Interaction and Virtual Reality Research Fields in Cognitive InfoCommunications”. In: *Applied Sciences* 11.6 (2021), p. 2646. DOI: 10.3390/app11062646. URL: https://doi.org/10.3390/app11062646.

[44] Soren Lauesen. *Software Requirements: Styles and Techniques*. London, UK: Pearson Education, 2002. ISBN: 0-201-74570-4.

[45] IEEE Computer Society. *IEEE Recommended Practice for Software Requirements Specifications*. Tech. rep. IEEE Std 830-1998. Available at: https://standards.ieee.org/standard/830-1998.html. Institute of Electrical and Electronics Engineers (IEEE), 1998, pp. 1–40. DOI: 10.1109/IEEESTD.1998.88286.

[46] Wei Liu, Tobias Placke, and K.T. Chau. "Overview of batteries and battery management for electric vehicles". In: *eTransportation* 12 (2022), p. 100166. DOI: 10.1016/j.etran.2022.100166. URL: https://www.sciencedirect.com/science/article/pii/S2352484722005716.

[47] Aslihan Örüm Aydin et al. "Lithium-Ion Battery Manufacturing: Industrial View on Processing Challenges, Possible Solutions and Recent Advances". In: *Batteries* 9 (2023), p. 555. DOI: 10.3390/batteries9110555. URL: https://doi.org/10.3390/batteries9110555.

[48] Jingyuan Zhao et al. "Battery safety: Fault diagnosis from laboratory to real world". In: *Journal of Power Sources* 598 (2024), p. 234111. DOI: 10.1016/j.jpowsour.2024.234111. URL: https://www.researchgate.net/profile/Jingyuan-Zhao-8/publication/378822364_Battery_safety_Fault_diagnosis_from_laboratory_to_real_world/links/65eb2a43b1906066b27662a1/Battery-safety-Fault-diagnosis-from-laboratory-to-real-world.pdf.

[49] American Chemical Society. *Hydrogen Fluoride - Molecule of the Week*. Accessed: 2025-04-27. 2025. URL: https://www.acs.org/molecule-of-the-week/archive/h/hydrogen-fluoride.html.

[50] Michael Burke et al. "Relative effectiveness of worker safety and health training methods". In: *American Journal of Public Health* 96.2 (2006), pp. 315–324. DOI: 10.2105/AJPH.2004.059840.

[51] Michael Orey. *Emerging Perspectives on Learning, Teaching, and Technology*. Ed. by Marisa Drexel. Licensed under a Creative Commons Attribution 3.0 License. The Global Text Project, 2010. URL: https://www.researchgate.net/publication/302947319_Emerging_Perspectives_on_Learning_Teaching_and_Technology.

[52] Klaas-Jan Stol and Brian Fitzgerald. "The ABC of Software Engineering Research". In: *ACM Transactions on Software Engineering and Methodology* 1.1 (2018), pp. 1–51. DOI: 10.1145/3241743. URL: https://doi.org/10.1145/3241743.

[53] Google Scholar. *Google Scholar - Search Engine for Scholarly Articles*. Accessed: 2024-02-20. 2024. URL: https://scholar.google.com/.

[54] IEEE Xplore Digital Library. *IEEE Xplore - Research Database for Engineering and Technology*. Accessed: 2024-02-20. 2024. URL: https://ieeexplore.ieee.org/.

[55] Henrik Söderlund. *The Battery Virtual Training (BVT) System: VR-Based Safety Training for Battery Assembly*. Unpublished presentation, GTC, Chalmers University of Technology, Sweden. 2025.

[56] Barbara A. Kitchenham et al. "Preliminary Guidelines for Empirical Research in Software Engineering". In: *IEEE Transactions on Software Engineering* 28.8

(2002), pp. 721–734. DOI: 10.1109/TSE.2002.1027796. URL: https://doi.org/10.1109/TSE.2002.1027796.

[57] Virginia Braun and Victoria Clarke. “Thematic Analysis”. In: *APA Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological*. Ed. by Harris Cooper et al. American Psychological Association, 2012, pp. 57–71. DOI: 10.1037/13620-004. URL: https://doi.org/10.1037/13620-004.

[58] Sandra G. Hart. “NASA-TASK LOAD INDEX (NASA-TLX); 20 Years Later”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.9 (2006), pp. 904–908. DOI: 10.1177/154193120605000909.

[59] Susana Rubio et al. “Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods”. In: *Applied Psychology* 53.1 (2004), pp. 61–86. DOI: 10.1111/j.1464-0597.2004.00161.x.

[60] Keith Vertanen. *NASA-TLX in HTML and JavaScript*. https://www.keithv.com/software/nasatlx/nasatlx.html. Accessed: 2025-05-14.

[61] Henrik Söderlund. *PhD Student at Chalmers University of Technology, Division of Production Systems*. Research focus on extended reality (XR) for upskilling and training in manufacturing industries. 2025. URL: https://research.chalmers.se/person/hensode.

[62] Zenodo. *Zenodo: Research. Shared.* https://zenodo.org/. Zenodo is an open-access repository developed under the European OpenAIRE program and operated by CERN. Accessed: 2025-05-12. 2013.

[63] Unity Technologies. *Meta XR - All-in-One SDK*. Accessed: 2025-04-26. 2025. URL: https://assetstore.unity.com/packages/tools/integration/meta-xr-all-in-one-sdk-269657.

[64] SketchUp. *3D Warehouse*. Accessed: 2025-04-26. 2025. URL: https://3dwarehouse.sketchup.com/.

[65] Natural Readers. *Natural Readers Online*. Accessed: 2025-04-26. 2025. URL: https://www.naturalreaders.com/online/.

[66] Meta Platforms Inc. *Meta Horizon Developer Documentation*. https://developers.meta.com. Accessed: 2025-05-08. 2024.

[67] Meta Platforms Inc. *Meta Quest 3 – Technical specifications*. https://www.meta.com/se/en/quest/quest-3/. Accessed: 2025-05-08. 2025.

# A

# Appendix

## A.1 Interview Questions

This section lists the semi-structured interview questions used in Phase 1.

1. Can you describe your experience using the VR training system? What stood out to you the most?
2. Do you feel that VR helped you understand and remember battery safety procedures better than other training methods? Why or why not?
3. Did the VR training feel realistic and relevant to your work? Is there anything missing that traditional training captures better than VR?
4. How easy or difficult was it to navigate and interact with the VR system? Did you encounter any issues?
5. If you could improve one thing about this VR training, what would it be? Would you recommend it to others?

## A.2 Survey Questions for Phase 1

This section lists all survey questions used in Phase 1.

### A.2.1 Background Information

These questions established demographic and contextual data about participants' prior experience with batteries, training methods, and VR.

1. **How old are you?**
   - Under 18 years old
   - 18-25 years old
   - 26-35 years old
   - 36-50 years old
   - 50+ years old
2. **Have you worked with battery management before?**
   - Yes
   - No
3. **What is your current role at work?** (Open-ended response)
4. **Have you received previous training in battery safety?**
   - Yes, online training
   - Yes, in a classroom
   - Yes, in a factory

- No
- Other: (Open-ended)

5. **Have you used Virtual Reality (VR) before?**
   - Yes, for games
   - Yes, for work training
   - Yes, for personal use
   - No
   - Other: (Open-ended)
6. **How comfortable are you with VR technology?** (Scale: 1 = Not at all comfortable, 5 = Very comfortable)

### A.2.2 Usability & User Experience

This section includes questions on system intuitiveness, comfort, feedback, and immersion.

7. **How intuitive did you find the VR training interface?** (Likert Scale: 1 = Not intuitive at all, 5 = Very intuitive)
8. **Did you experience any discomfort while using the VR system?**
   - Motion sickness
   - Eye strain
   - Cognitive overload
   - Other: (Open-ended)
9. **Did the system provide clear and immediate feedback on your actions?**
   - Yes
   - No
   - Other: (Open-ended)
10. **How educational did you find VR training compared to previous learning methods you have experienced?** (Scale: 1 = Not educational at all, 5 = Very educational)
11. **How clear were the instructions during the VR training? (How to navigate, perform tasks, and understand the goals of the training, etc.)** (Scale: 1 = Very unclear, 5 = Very clear)
12. **How did you experience the immersion in the VR training? Did it feel real and engaging?** (Scale: 1 = Not engaging at all, 5 = Very engaging and realistic)
13. **How realistic did the incidents you experienced in the training feel? (e.g., electric shocks, battery handling, risky situations)** (Scale: 1 = Not at all realistic, 5 = Very realistic)

### A.2.3 Effectiveness of Learning & Knowledge Retention

These questions measured participants' confidence, knowledge gain, and memory of safety procedures after training.

14. **After completing the VR training, how confident do you feel in handling battery safety procedures?** (Scale: 1 = Not confident at all, 5 = Very confident)
15. **To what extent has the VR training helped you understand battery safety risks? (e.g., handling chemicals, exposure to high voltage)** (Scale: 1 = No improvement at all, 5 = Very big improvement in understanding)
16. **How well can you recall key safety procedures after using the VR system?** (Scale: 1 = I don't remember anything, 5 = Remember very well)
17. **To what extent did you feel that the VR training made you reflect on safety incidents and how you would handle them in real life?** (Scale: 1 = No reflection at all, 5 = Very deep reflection and insight)

### A.2.4 Technical & Design Improvements

Open-ended questions to suggest improvements or report technical difficulties.

18. **Were there any technical difficulties you encountered while using the VR system? If "Yes", explain below; if not, write "No"** (Open-ended response)
19. **What improvements would you suggest to make VR training more effective? Is there any safety scenario you feel is missing?** (Open-ended response)
20. **Is there any additional feedback/opinion/comment you would like to add? It can be both positive and negative!** (Open-ended response)

## A.3 Interview Questions for BCG Trainers

This section lists the questions used in the interview with the BCG trainers.

1. Can you describe the typical workflow in the "raw to cell" process at BCG?
2. What are the main safety risks and critical control points in this part of the process?
3. In your experience, what parts of the process are hardest to teach in traditional (non-VR) training?
4. Are there any steps in the process that are difficult to simulate safely in the real world?
5. Where do you see the biggest opportunity for VR to enhance or complement current training methods?
6. Based on your experience, what improvements would you suggest for making the VR training more realistic or effective?
7. What emergency situations or incorrect actions would be valuable to include in the VR training?
8. How do you see the role of VR evolving in battery manufacturing training at BCG?

## A.4 Detailed results from the Usability Survey

### A.4.1 Phase 1 detailed quantitative responses

Mean and standard deviation per theme and participant group calcu- lated from the Likert-scale responses in the survey from Phase 1.

| Theme | BAF (Mean ± SD) | HVE (Mean ± SD) |
|---|---|---|
| VR Comfort | 3.65 ± 1.04 | 4.13 ± 0.94 |
| Interface Intuitiveness | 4.23 ± 0.78 | 4.23 ± 0.77 |
| Immersion | 4.04 ± 0.92 | 4.13 ± 0.73 |
| Incident Realism | 3.69 ± 1.11 | 3.50 ± 1.33 |
| Confidence | 3.81 ± 0.91 | 3.57 ± 0.90 |
| Understanding Risks | 3.69 ± 0.90 | 3.90 ± 0.80 |
| Safety Recall | 4.08 ± 0.82 | 4.07 ± 0.94 |
| Reflection | 3.94 ± 0.91 | 3.97 ± 0.76 |
| Instruction Clarity | 4.06 ± 0.93 | 4.13 ± 0.94 |
| Educational Value | 4.27 ± 0.84 | 4.20 ± 0.66 |

**Table A.1:** Mean and standard deviation per theme and participant group calculated from the Likert-scale responses in the survey.

### A.4.2 Phase 4 detailed quantitative responses

Mean and standard deviation per theme calculated from the Likert-scale responses in the survey from Phase 4.

| Theme | Mean | SD |
|---|---|---|
| VR Comfort | 4.25 | 0.62 |
| Interface Intuitiveness | 4.25 | 0.87 |
| Immersion | 4.33 | 0.65 |
| Incident Realism | 3.42 | 0.79 |
| Confidence | 2.92 | 1.00 |
| Understanding Risks | 3.33 | 0.98 |
| Safety Recall | 4.33 | 0.89 |
| Reflection | 3.83 | 0.58 |
| Instruction Clarity | 3.75 | 0.97 |
| Educational Value | 4.00 | 0.43 |

**Table A.2:** Survey results summary from Phase 4: Mean and standard deviation per theme

## A.5 Identified Codes from Thematic Analysis

This section presents the thematic coding results from the qualitative data collected during the study. Codes were derived from open-ended survey responses and group interviews and organized into themes relevant to usability, realism, instruction, and system improvement.

### A.5.1 Phase 1: Analysis of Open-ended Survey Responses

| Code | Theme | Frequency |
|---|---|---|
| Object placement difficulties | Interaction and usability | 6 |
| Inability to grab or drop tools correctly | Interaction and usability | 5 |
| Lag or system delay | Technical issues | 7 |
| Multimeter malfunction | Technical issues | 3 |
| Environment shifted unexpectedly | Technical issues | 4 |
| VR physics inaccuracies | Realism and immersion | 2 |
| Suggestions for more scenarios | Content and scenario improvement | 6 |
| Desire for clearer instructions | Instruction clarity | 4 |
| Request for more feedback on task performance | Instruction clarity | 3 |
| General positive feedback | Overall impression | 5 |
| Desire for collaborative/multi-user training | Collaboration and multi-user | 3 |
| Discomfort or eye strain | Physical discomfort | 2 |

**Table A.3:** Identified codes, their frequencies, and corresponding themes from open-ended survey responses.

### A.5.2 Phase 1: Analysis of Group Interviews

| Code | Theme | Frequency |
|---|---|---|
| Wanted to do tasks with others | Collaboration and multi-user | 5 |
| Better than traditional classroom training | Comparison to traditional training | 6 |
| Missed instructions | Instruction clarity | 9 |
| Unclear steps | Instruction clarity | 6 |
| Difficult to understand what to do | Instruction clarity | 7 |
| Felt realistic / like real life | Realism and immersion | 11 |
| Immersive experience | Realism and immersion | 7 |
| Jumped from sound / surprising elements | Shock and consequence | 8 |
| Shock effect helped learning | Shock and consequence | 7 |
| Practical exercise helped memory | Learning by doing | 9 |
| Being able to try myself was good | Learning by doing | 10 |
| Suggested adding more scenarios | Suggestions for improvement | 7 |
| Eye strain and other discomforts | Physical discomfort | 6 |
| System bugs / technical issues | Technical issues | 7 |
| Easier to remember after VR | Retention and memory | 5 |

**Table A.4:** Identified codes, their frequencies, and corresponding themes from group interview analysis.

### A.5.3 Phase 4: Analysis of Open-ended Survey Responses

| Code | Theme | Frequency |
|---|---|---|
| Request for more tasks/scenarios | Scenario length and variety | 3 |
| Scenario could be longer | Scenario length and variety | 2 |
| Add more safety scenarios (e.g., fire, spill) | Scenario length and variety | 1 |
| Environment not fully realistic | Realism and environment | 1 |
| Mismatch between instruction and environment | Realism and environment | 1 |
| Instructions not always visible | Instructional clarity and timing | 2 |
| Need clearer or embedded instructions | Instructional clarity and timing | 1 |
| Suggestion for audio prompts | Instructional clarity and timing | 1 |
| Dropped object or interaction glitch | Technical and behavioral issues | 1 |
| Positive: better than traditional methods | General impressions | 2 |
| Positive: engaging and clear | General impressions | 2 |

**Table A.5:** Identified codes and their frequencies from Phase 4 open-ended survey responses.

## A.6 Requirement Mapping

This section provides a comprehensive mapping of all requirements to their corresponding themes and data sources. This table provides full traceability from data collection to specification formulation. Some requirements appear under multiple themes where applicable, to reflect their relevance.

| **Theme** | **Data Source** | **Mapped Requirements** |
|---|---|---|
| VR Comfort | Quantitative Survey | NFR-1, NFR-2, NFR-11, UR-1, UR-8, FR-2, FR-13 |
| Interface Intuitiveness | Quantitative Survey | FR-4, UR-2, UR-9, FR-11 |
| Immersion | Quantitative Survey | FR-1, FR-12, NFR-5, UR-3, NFR-12 |
| Incident Realism | Quantitative Survey, Group Interviews | FR-B1, FR-B2, FR-B5 |
| Confidence | Quantitative Survey, Group Interviews | FR-10, UR-6, FR-8 |
| Understanding Risks | Quantitative Survey, Literature Review, Stakeholder Interview | FR-B1, FR-B3, UR-6 |
| Safety Recall | Quantitative Survey | FR-8, DR-1, UR-5 |
| Reflection | Quantitative Survey, Qualitative Survey | UR-4, FR-2, FR-8, DR-4, FR-3 |
| Instruction Clarity | Quantitative Survey, Group Interviews | UR-2, UR-9, FR-4, FR-5, FR-9, UR-10, NFR-4, NFR-6 |
| Educational Value | Quantitative Survey, Literature Review | FR-7, FR-10, UR-6, DR-1, FR-B8 |
| Realism and Engagement | Qualitative Survey, Group Interviews | FR-B1, FR-B2, FR-12 |
| Learning and Knowledge Retention | Qualitative Survey, Group Interviews, Literature Review | FR-10, DR-1, UR-6, FR-3, DR-6, UR-4, FR-B8 |
| Technical Issues | Qualitative Survey, Group Interviews, User Testing Observation | NFR-1, NFR-11, SR-7 |
| Guidance and Feedback | Qualitative Survey, Group Interviews | FR-5, UR-2, FR-8, NFR-6 |
| Suggested Improvements | Qualitative Survey, Group Interviews | NFR-3, NFR-5, UR-9 |
| Hardware | User Testing, Best Practice | SR-8 |

| Theme | Data Source | Mapped Requirements |
|---|---|---|
| Applicability to Real Work | Qualitative Survey, Group Interviews | UR-6, FR-B3, FR-B5, FR-B4 |
| Emergency Readiness | Stakeholder Interview | FR-B5, FR-6, UR-6 |
| Contamination Risk | Stakeholder Interview | FR-B6, FR-B7 |
| Cognitive Load | Literature Review | NFR-6, UR-3 |
| User-Centered Design | Literature Review | UR-2, UR-9, NFR-3 |
| Data Privacy | Literature Review | DR-2, DR-5, SR-5 |
| Motion Sickness Avoidance | Literature Review | NFR-8, NFR-12, UR-7 |
| System Performance | User Testing Observation, Best Practice | NFR-7, SR-1, SR-2, SR-4, SR-6, SR-7, SR-10, SR-11 |
| Loading and Feedback | User Testing Observation | NFR-8, NFR-9, NFR-10, SR-3 |
| Calibration and Tracking | User Testing Observation | SR-4, UR-11 |
| Usability Benchmarks | Best Practice / Standards | NFR-3, FR-3, DR-3 |
| GDPR Compliance | Best Practice / Standards | SR-5, DR-2, DR-5 |
| Modularity and Maintenance | Best Practice / Standards, Qualitative Survey | SR-9, FR-13, NFR-8 |

**Table A.6:** Themes and data sources mapped to the corresponding requirements.

## A.7 PoC Scenario Code

This section includes two Unity C# scripts used in the implementation of the VR training scenario. The scripts define how the scenario responds to user actions such as dropping objects, triggering events, and completing tasks.

### A.7.1 CellDrop.cs Code

Handles task logic, feedback progression, and user interaction for the battery spill and evacuation scenario.

```
using UnityEngine;
using System.Collections;
using TMPro;
using System;
using UnityEngine.UI;
using Oculus.Interaction.HandGrab;

public class CellDrop : MonoBehaviour
{
    // User info
    const string glyphs = "
       abcdefghijklmnopqrstuvwxyz0123456789";
    public int userNameLength = 6;
    private string myString;

    public TMP_Text userNameTextObj;
    public TMP_Text dateTextObj;
    public TMP_Text progressTextObj;

    // Progress tracking
    private int tasksCompleted = 0;

    // Tick marks on whiteboard
    public TMP_Text tickMark1;
    public TMP_Text tickMark2;
    public TMP_Text tickMark3;
    public TMP_Text tickMark4;

    // Scenario logic
    public bool dropped;

    public GameObject GameController;
    public ParticleSystem[] PS;
    public GameObject GrabObject;
    public AudioSource DropSound;

    public GameObject canvasPickup;
    public GameObject canvasCheck;
```

```
public AudioSource InstructionAudio;
public AudioSource EvacuationAudio;

public GameObject EmergencyExitSign;
public GameObject EmergencyExitZone;

public GameObject FaceMaskObject;
public GameObject FaceMaskObjectOn;

public GameObject BlackOut;
public GameObject completionTextObj;
public GameObject ImageBlackOut;


void Start()
{
    GenerateAndDisplayUserInfo();
}

/*
Generates the user information displayed on the
   whiteboard. This includes an anonymous
username, today's date and the progress shown in
   percentages.
*/
void GenerateAndDisplayUserInfo()
{
    if (StaticDataGTC.UserName == null)
    {
        myString = "";
        for (int i = 0; i < userNameLength; i++)
        {
            myString += glyphs[UnityEngine.Random.Range
               (0, glyphs.Length)];
        }

        StaticDataGTC.UserName = myString;
        userNameTextObj?.SetText("User Name: " + myString
           );
    }
    else
    {
        myString = StaticDataGTC.UserName;
        userNameTextObj?.SetText("User Name: " + myString
           );
    }

    string today = DateTime.Today.ToString("yyyy-MM-dd");
```

```
        dateTextObj?.SetText("Date: " + today);
        StaticDataGTC.Date = today;

        progressTextObj?.SetText("Progress: 0%");
    }

    /*
    Handles the first step of the scenario, which is to drop
       the battery cell on the floor.
    When dropped, the first step shows "completed!" on the
       whiteboard, then the current canvas
    is hidden and the next canvas with information shows
    */
    void OnCollisionEnter(Collision col)
    {
        if (!dropped && col.gameObject.CompareTag("floor"))
        {
            dropped = true;

            if (tickMark1) tickMark1.text = "Completed!";
            UpdateProgress();

            GameController?.GetComponent<CellSmokeSequence>()
               ?.CellDropped();
            DropSound?.Play();

            if (canvasPickup != null)
            {
                canvasPickup.SetActive(false);
                Destroy(canvasPickup);
            }

            if (canvasCheck != null)
            {
                canvasCheck.SetActive(true);

            }

            StartCoroutine(PlaySmokeThenAudio());
        }
    }

    /*
    Plays smoke after a set delay, which is the second step
       of the scenario.
    Then plays the instruction audio after a set delay. After
        the audio is played, the mask is outlined and can
    now be grabbed and put on. Finally the second task on the
        whiteboard is marked complete.
```

```
*/
IEnumerator PlaySmokeThenAudio()
{
    float delayBetweenSmokes = 2f;

    for (int i = 0; i < PS.Length; i++)
    {
        if (PS[i] != null)
        {
            PS[i].Play();
            yield return new WaitForSeconds(
               delayBetweenSmokes);
        }
    }

    yield return new WaitForSeconds(17f);

    if (InstructionAudio != null)
    {
        InstructionAudio.Play();
    }

    if (FaceMaskObject != null)
    {
        Outline outline = FaceMaskObject.GetComponent<
           Outline>();
        if (outline != null)
        {
            outline.enabled = true;
        }

        FaceMaskObject.transform.Find("
           ISDK_HandGrabInteraction").GetComponent<Oculus
           .Interaction.HandGrab.HandGrabInteractable>().
           enabled = true;

    }

    if (tickMark2) tickMark2.text = "Completed!";
    UpdateProgress();

    StartCoroutine(WaitForMaskAndEvacuate());
}

/*
When the user has put on the face mask, task 3 is marked
   complete on the whiteboard,
and the evacuation audio is played. Then the emergency
   exit is outlined and enabled.
```

```
*/
IEnumerator WaitForMaskAndEvacuate()
{
    while (FaceMaskObject != null && FaceMaskObject.
       activeInHierarchy)
    {
        yield return null;
    }

    if (tickMark3) tickMark3.text = "Completed!";
    UpdateProgress();

    if (EvacuationAudio != null)
    {
        EvacuationAudio.Play();
    }

    if (EmergencyExitSign != null)
    {
        EmergencyExitSign.SetActive(true);
        Renderer renderer = EmergencyExitSign.
           GetComponent<Renderer>();
        if (renderer != null)
        {
            renderer.material.EnableKeyword("_EMISSION");
        }
    }

    if (EmergencyExitZone != null)
    {
        EmergencyExitZone.SetActive(true);
    }
}

// Function to update the progress percentage
void UpdateProgress()
{
    tasksCompleted++;
    float progressPercent = (tasksCompleted / 4f) * 100f;
    progressTextObj?.SetText("Progress: " +
       progressPercent + "%");
}

public void Evacuated()
{
    StartCoroutine(Fade());
}

/*
```

```
    Evacuation logic. When the user has entered the outlined
       emergency exit zone,
    a white fading screen (here BlackOut) appears. Then a
       completion text appears and
    the face mask object is hidden (to not show through the
       white screen).
    */
    private IEnumerator Fade()
    {
        float transparency = 0.0f;
        float duration = 2.0f;

        float normalizedTime = 0;


        if (BlackOut != null)
        {
            BlackOut.SetActive(true);
        }


        while (normalizedTime <= 1f)
        {
            normalizedTime += Time.deltaTime / duration;

            transparency = transparency + 0.01f;
            ImageBlackOut.GetComponent<Image>().color = new
               Color(255, 255, 255, transparency);

            yield return null;
        }


        if (completionTextObj != null)
        {
            completionTextObj.SetActive(true);
        }

        FaceMaskObjectOn.SetActive(false);
    }
}
```

### A.7.2 Head.cs Code

Detects user entry into the emergency exit zone and triggers scenario completion.

```
using UnityEngine;

public class Head : MonoBehaviour
```

```
{
    public CellDrop cellDropScript;

    void OnTriggerEnter(Collider other)
    {
        if (other.CompareTag("EmergencyExit"))
        {
            if (cellDropScript != null)
            {
                cellDropScript.Evacuated();
            }
        }
    }
}
```

## A.8 Requirement Fulfillment Overview by Category (Phase 3)

This section summarizes how many requirements were addressed in Phase 3 and how many were no addressed.

| **Requirement Category** | **Total** | **Not Addressed** | **Total Addressed** |
|---|---|---|---|
| FR | 13 | 6 | 7 |
| FR-B | 8 | 2 | 6 |
| UR | 11 | 3 | 8 |
| NFR | 12 | 7 | 5 |
| SR | 11 | 1 | 10 |
| DR | 6 | 5 | 1 |
| **Total** | **61** | **24** | **37** |

**Table A.7:** Requirement fulfillment overview by category and in total, showing the number of total, addressed, and not addressed requirements.

## A.9 Requirement Fulfillment Source Matrix (Phase 3)

This section summarizes how requirements were addressed in the project, whether through hardware, implemented features in the PoC, or planned verification in Phase 4.

| Requirement ID | Meta Quest 3 Hardware | Implemented in PoC (Phase 3) | Needs Phase 4 Validation |
|---|---|---|---|
| FR-1 | | ✓ | |
| FR-2 | | ✓ | |
| FR-5 | | ✓ | |
| FR-6 | | ✓ | ✓ |
| FR-7 | ✓ | ✓ | ✓ |
| FR-9 | | ✓ | |
| FR-12 | | ✓ | ✓ |
| FR-B1 | | ✓ | |
| FR-B2 | | ✓ | |
| FR-B3 | | ✓ | |
| FR-B5 | | ✓ | |
| FR-B7 | | ✓ | |
| FR-B8 | | ✓ | |
| UR-1 | | ✓ | ✓ |
| UR-2 | | ✓ | |
| UR-3 | | ✓ | ✓ |
| UR-4 | | ✓ | |
| UR-5 | | ✓ | |
| UR-6 | | ✓ | ✓ |
| UR-7 | | ✓ | ✓ |
| UR-11 | ✓ | ✓ | ✓ |
| NFR-4 | | ✓ | |
| NFR-6 | | ✓ | ✓ |
| NFR-7 | | ✓ | ✓ |
| NFR-9 | | ✓ | |
| NFR-12 | ✓ | ✓ | |
| SR-1 | | ✓ | |
| SR-2 | ✓ | | |
| SR-3 | ✓ | | |
| SR-4 | ✓ | ✓ | ✓ |
| SR-5 | | ✓ | ✓ |
| SR-6 | | ✓ | ✓ |

| Requirement ID | Meta Quest 3 Hardware | Implemented in PoC (Phase 3) | Needs Phase 4 Validation |
|---|---|---|---|
| SR-7 | ✓ | ✓ | ✓ |
| SR-8 | ✓ | | |
| SR-10 | ✓ | ✓ | |
| SR-11 | ✓ | ✓ | ✓ |
| DR-2 | ✓ | ✓ | ✓ |

**Table A.8:** Matrix showing how each requirement was fulfilled and what needed to be validated in Phase 4.

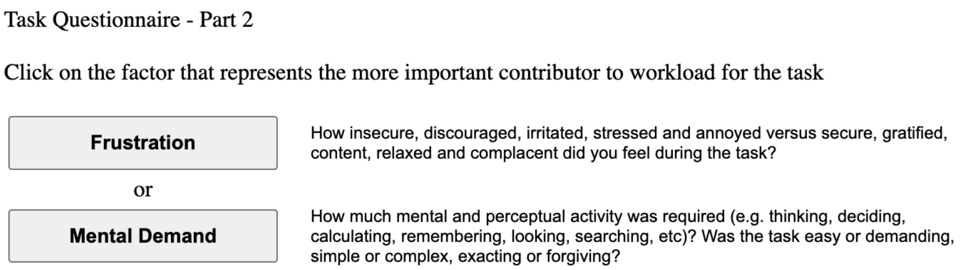
## A.10 Additional Survey Questions for Phase 4

This section lists additional survey questions used in Phase 4 to evaluate the newly implemented features of the PoC system, such as the help and reset buttons.

21. **Did you use the "Help" button during the VR scenario?** (Options: Yes / No)
22. **If you used the Help button, how helpful did you find it?** (Scale: 1 = Not helpful at all, 5 = Very helpful)
23. **Do you think the Help button would be important in a more complex training system with multiple scenarios and tasks?** (Options: Yes / No / Other)
24. **Did you use the "Restart" button to restart the scenario during your training?** (Options: Yes / No)
25. **If you used the Restart button, how many times did you use it?** (Options: Once / Twice / More than twice)
26. **Did you feel like the scenario guided you clearly enough to respond correctly during the emergency moment (gas leak and evacuation)?** (Scale: 1 = Very unclear, 5 = Very clear)

## A.11 NASA Task Load Index Survey

This section shows two examples of questions in the NASA-TLX survey, used to measure perceived workload across six dimensions. These questions were used in both Phase 1 and Phase 4.



Task Questionnaire - Part 2

Click on the factor that represents the more important contributor to workload for the task

**Frustration**

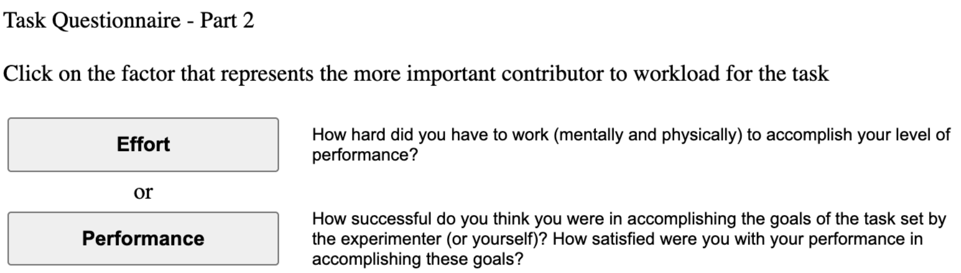
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

or

**Mental Demand**

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

**Figure A.1:** Example of pairwise comparison question in the NASA-TLX workload survey.



Task Questionnaire - Part 2

Click on the factor that represents the more important contributor to workload for the task

**Effort**

How hard did you have to work (mentally and physically) to accomplish your level of performance?

or

**Performance**

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

**Figure A.2:** Another example of pairwise comparison question in the NASA-TLX workload survey.

## A.12 Protocol for Measured Real Task

This section contains the protocol used for observing and measuring the Phase 4 real-world evaluation task. Tables include action completion data, time-stamps and notes.

| Participant ID | Battery dropped? | Inspected battery? |
|---|---|---|
| CG-P1 | Yes | Yes |
| CG-P2 | Yes | Yes, but did not pick it up |
| TG-P1 | Yes | Yes |
| TG-P2 | Yes | Yes |
| TG-P3 | Yes | Yes, inspected on the floor |
| CG-P3 | Yes | Yes |
| CG-P4 | Yes | No, barely |
| TG-P4 | Yes | No |
| TG-P5 | Yes | Yes |
| TG-P6 | Yes | Yes |
| TG-P7 | Yes | Yes |
| CG-P5 | Yes | No, barely |
| CG-P6 | Yes | No, barely |
| CG-P7 | Yes | Yes |

**Table A.9:** Participant actions before gas leak audio.

| Participant ID | Gas mask used + time | Reacted to alarm + time | Evacuated correctly + time | Total time |
|---|---|---|---|---|
| CG-P1 | Yes, 01:20 | Yes, 01:32 | Yes, 01:50 | 02:02 |
| CG-P2 | Yes, before gas leak 01:17 | Yes, 01:39 | Yes, 01:43 | 01:50 |
| TG-P1 | Yes, 00:38 | Yes, 00:49 | Yes, 00:57 | 01:09 |
| TG-P2 | Yes, 00:49 | Yes, 00:57 | Yes, 01:01 | 01:03 |
| TG-P3 | Yes, 00:30 | Yes, 00:33 | Yes, 00:40 | 00:45 |
| CG-P3 | Yes, 00:38 | Yes, 01:20 | Yes, 01:21 | 01:28 |
| CG-P4 | Yes, 00:32 | No | No | 01:56 |
| TG-P4 | Yes, 00:31 | Yes, 00:40 | Yes, 00:45 | 00:51 |
| TG-P5 | Yes, 00:28 | Yes, 00:50 | Yes, 00:56 | 00:59 |
| TG-P6 | Yes, 00:41 | Yes, 00:56 | Yes, 00:58 | 01:01 |
| TG-P7 | Yes, 01:04 | Yes, 01:36 | Yes, 01:37 | 01:40 |
| CG-P5 | Yes, 00:18 | Yes, 00:50 | Yes, 00:52 | 00:59 |
| CG-P6 | Yes, 00:36 | Yes, 01:17 | Yes, 01:20 | 01:23 |
| CG-P7 | Yes, 02:26 | Yes, 02:47 | Yes, 02:55 | 02:59 |

**Table A.10:** Participant response timing and task completion.

| Participant ID | Comments |
|---|---|
| CG-P1 | Used surgical mask before gas mask |
| CG-P2 | Wanted to put on surgical mask but changed their mind |
| TG-P1 | – |
| TG-P2 | Used the face mask before warning but after gas sound |
| TG-P3 | Used surgical mask before gas mask |
| CG-P3 | – |
| CG-P4 | Did not evacuate |
| TG-P4 | – |
| TG-P5 | – |
| TG-P7 | Took off their real jewelry, put on surgical mask |
| CG-P7 | Took their time inspecting the table. Put on surgical mask. |

**Table A.11:** Additional notes from the measured real task.