KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Virtual Reality Serious Game for Chemical Lab Safety Training in Industry and Academia:

from Design, Development to Increased Motivation and Engagement

Philippe Chan

Supervisors:
Prof. dr. ir. Kristel Bernaerts
Dr. Jean-Luc Dubois

Co-supervisor:
Prof. dr. ir. Tom Van Gerven

Dissertation presented in partial fulfilment
of the requirements for the degree of
Doctor of Engineering Science (PhD):
Chemical Engineering

March 2023

Virtual Reality Serious Game for Chemical Lab Safety Training in Industry and Academia:

from Design, Development to Increased Motivation and Engagement

Philippe CHAN

Examination committee:
Prof. dr. ir. Dirk Vandermeulen, chair
Prof. dr. ir. Kristel Bernaerts, supervisor
Dr. Jean-Luc Dubois, supervisor (Arkema)
Prof. dr. ir. Tom Van Gerven, co-supervisor
Prof. dr. Mario Smet
Prof. dr. Peter De Graef (Thomas Moore)
Prof. dr. Karin Coninx (UHasselt)
Ing. Hans Musters (ACTA vzw)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Engineering Science (PhD): Chemical Engineering

March 2023

This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement 812716

Uitgegeven in eigen beheer, Philippe Chan, Celestijnenlaan 200F box 2424, B-3001 Leuven (Belgium)

Preface

> *"Do not pray for an easy life,*
> *pray for the strength to endure a difficult one."*
>
> Bruce Lee, 1940-1973

Finally... I have accomplished something I have been longing for for a long time. Ever since I was a little kid, I always wanted to achieve great things later in my life. Becoming a chemical scientist and getting a doctoral degree were some of the goals that I have set to achieve. I knew from the beginning that these goals were difficult to get, but I also knew that the most important part is the journey towards it. It is by constantly challenging myself during this adventure that I gain incredibly useful experience, and at each step forward, I have become a better person than before. However, I must say that this would not have been possible without the support of all the people close to me. As well as the people I've met along the way who have become people I cherish a lot, whether it be from a random encounter at the bar, at the climbing gym or even in the air plane. For this reason, I want to express my gratitude to these people in this section of my dissertation.

First and foremost, I would like to thank the people who have been mentoring me in this doctoral research. Dr. Jean-Luc Dubois has taught me how to be more professional in a chemical company environment. I really admire his exceptional experience in several big research projects, including many EU projects. Prof. Kristel Bernaerts has helped me many times by giving very useful feedback on manuscripts and presentations. Every time when I could not see how to simplify my explanation, she provided me insights from another perspective that help me get back on track. Our minds were sometimes in sync. Prof. Tom Van Gerven is officially my 'co-'supervisor, but to me, this 'co-' can be dropped because he guided me like one of his own Ph.D. student. Many times, he was even the first to respond to my questions and manuscripts. These supervisors have brought great support and guidance in my growth and development as a researcher.

I would also like to express my sincere gratitude to the members of my Examination Committee for taking the time and effort to read my doctoral dissertation and make insightful comments during the preliminary defence. I appreciate your unwavering support and I have learned a lot from our conversations.

In the three years in Lyon in France, I am very grateful I had such welcoming and supportive colleagues at Arkema. Despite they were managers in chemical R&D and had no clue about what I did, they still supported me with this project and also with the life and culture of working in France. From day one to last, we only spoke in French in our coffee and lunch breaks, which helped me to be so much more confident in my French speaking skills. Merci du fond du cœur Jérôme, Anne, Laurent, Dominique, David et Gerard!

Also in an international sense, I am forever grateful for being a part of the CHARMING family. When I encountered the application to become a doctoral researcher for the ETN-CHARMING project, I thought that this is my destiny, as if it is meant to be. The combination of chemistry, games, VR/AR and educational sciences is the perfect match of who I deeply am. On top of that, all of us charming colleagues had an unusual great synergy together, as if we were already friends for many years. We were all facing the same problem in our projects: it is too interdisciplinary. Especially social sciences are totally different from natural sciences... Somehow, we still managed to learn from each other and have great times during Network Wide Events. So, many thanks to Michaela, Yuanyuan, Tim, Chioma, Ryo, Yusra, Silvia and Nina. Moreover, a special shout-out to fellow KU Leuven ESRs: Michael, Pedro, Serkan, Jessica, Sanne. I really missed hanging out with fellow ESRs when I was in France. In particular, I want to thank Sofia, because she had the German version of my journey: started almost at the same time; both working in a chemical company in another country, but a Ph.D. from KU Leuven; both supervised by Kristel. We helped each other countless of times — the same goes for the whole CHARMING consortium. In addition, special thanks to Tom and Rabab for managing such extraordinary project.

Since I arrived in Leuven, I would also gladly thank fellow researchers/staff at KU Leuven Chem&Tech. The people from CREaS, ProCESS and SMaRT have shown me how important it is to be surrounded by fellow colleagues with the same struggle of doing research and to help each other out when needed. We are all in this together. Special appreciations to Kenneth, Justien and Luan who are great office mates and helped me a lot to get acquainted with the KU Leuven system. Also, I really enjoyed the board game evenings with Jonas, Pieter A., Glen, Pieter D.W., Michiel and Holger. I wish them good luck with their Ph.D. journey as well.

Of course, this journey wouldn't be the same without the awesome memories with close friends! First, I am grateful for Michael and Pedro to let me join their friends group when I arrived in Leuven. We spend so many fun times together that they have become my close friends as well. Thanks for all the beers and laughs: Corentin, Martijn, Senne, Dionysia, Keily, Kevin, Sasha, Rita, Homila and Jalil. Let's meet up for some drinks and games this week!

This goes without saying that I have met friends for life in Lyon as well. What began with drunk evenings turned into an unforgettable friendship bond: Jo, Laura, Wiam, Julie and Gladys. What began as random climbing session turned into climbing partners for eternity: Noza, Komodo, Jin and Max. We had such awesome memorable trips that I am forever thankful to have met these people. I'm ready to plan our next adventure!

Moreover, I am deeply grateful for my close friends from before this doctoral project, in particular: Ronald, Chiel, Cédric, Babs, Preben, Luca, Jasper, Arne and Bart. Without the help I had from them in my uni years, I wouldn't be able to even start this journey. They have always pushed me to go beyond my capabilities, and in fact, we encouraged each other to keep doing what we love. Here, I'm also talking about my friends closer to my home town: Andrei, Vlad, Radu, Thomas, Jerko, David, William and Yuri. The fun times we had together really helped me to relax from the intense periods. I will never forget all the great memories we had and for our long-standing friendship that will stand for an even longer time. Zeg, massa's bedankt é matjes!

Last but not least, I want to express my heartfelt gratitude to my family both in Belgium and in Hong Kong. They have been rooting for me since the very beginning. Although, I have been silent for a part of the family, they still have my utmost respect and can count on me at any time. We are one family. Also, special thanks to my two older brothers Chi and Wellen who have taken care of me since I was little. I have learned both good and bad things from them, but they have shaped me who I am to this day. Especially, my mom and dad have made me go beyond my own limits to achieve greater things. Because, I understand that they have been working extremely hard, just to be able to give us a fulfilling life. So, I don't go through this challenging journey just for myself, I also do it for them. 多謝晒大家!

There are a lot more people I would like to thank, but these pages cannot go to eternity unlike my gratitude to them.

My journey to greatness has not ended yet.
To Be Continued...

Abstract

Chemical laboratories are workplace environments where risks of injury and even fatality are always present. Laboratory workers, in academia and industry, are required to have proficient safety awareness in order to minimise risk to an acceptable level. However, investigation reports and reviews on laboratory accidents reveal that, due to lack of safety awareness and at-risk behaviour, unfortunate events still happen. Insufficient safety training is often identified as one of the causes for these accidents.

Commonly, safety training interventions are given by means of conventional teaching methods, such as classroom lectures, videos and printed manuals. However, with these methods, trainees are required to listen to the instructor or read long textual paragraphs. They are passive in their learning process and can lose their attention quickly when they are not motivated, thus making the safety training ineffective. Virtual reality (VR) technology can be used as a hands-on training tool to simulate dangerous situations in a virtual environment where trainees can train their safety awareness without causing real harm to themselves or others. A combination of this technology with a serious game approach could make safety training more motivating and engaging than conventional training methods.

Therefore, this doctoral research project exploits immersive virtual reality technology to fully immerse the learning in a virtual laboratory environment, where laboratory safety is trained without putting the learner in danger. In particular, this research wants to obtain a better understanding of the motivation and engagement of employees following safety training in an immersive VR environment. However, designing and developing a VR serious game for this specific purpose is a challenging task. There are many factors to consider for an optimal training game. For this reason, the aim of this research is to not only verify if motivation and engagement are increased by using this immersive tool, but also to formulate guidelines on how to design, develop and implement such VR games for safety training. In this dissertation, an example is shown of

the design and development process of a VR serious game, called VR LaboSafe Game. The guidelines established from this process are generic and can be applied to the development of other similar VR safety training programmes. These guidelines include the integration of instructional design, game design and VR considerations. Also, online resources and optimisation techniques are suggested in order to support self-directed learning of VR game development.

During the development of VR LaboSafe Game, tests were performed on academic and industrial population in order to analyse the system usability and simulator sickness after playing early versions of the game. This is to ensure that the game is easy-to-use and does not induce severe simulator sickness before performing further evaluation studies. A test with a first version of VR LaboSafe Game 1.0 revealed that displaying tutorial instructions only in a textual format could be the cause of low system usability. This test also showed that long continuous VR exposure (>40 minutes) could lead to mild simulator sickness symptoms. Therefore, the VR game was adjusted to reduce textual information, include spoken instructions and introduce guiding animations. Players are also allowed to take breaks from VR between game levels. A test with this updated version VR LaboSafe Game 2.0 indeed confirmed that system usability and simulator sickness are improved compared to the version 1.0. This VR LaboSafe Game 2.0 is then suitable to use in further studies.

In an evaluation study with the VR LaboSafe Game, more person-centred variables were examined in order to determine how motivated and engaged employees of a chemical company are for safety training with a conventional method and with VR serious game. From this experiment, lab technicians and managers tend to be personally motivated to follow safety training for their own safety and the safety of others. However, they find safety training to be repetitive, passive and easy to lose attention when it is presented with conventional training methods. Results show that employees find safety training with a VR serious game more intrinsically motivating and engaging than conventional training methods. However, complicated usability and the unfamiliarity of VR can affect their autonomous motivation for safety training. Notably, older employees (above 50 years old) have more difficulties with using VR headsets than younger employees (under 30 years old). Despite this complication, employees intend to follow VR safety training more than with a conventional method. It is suggested to combine conventional methods with VR as complementary tool and provide more frequent and smaller sessions, gradually introducing VR technology to beginners.

The outcomes of this thesis are two-fold: 1) a set of guidelines are demonstrated about the design, development and implementation of VR serious games for safety training, and 2) a functional artefact is created to teach lab safety awareness. With the use of this artefact, it has been proven that VR

serious games increases the intrinsic motivation and engagement of trainees for safety training. However, before fully deploying VR training programmes, consideration must be taken into account to let people familiarise with VR technology.

Korte inhoud

Chemische laboratoria zijn werkplaatsen waar risico's op verwondingen en zelfs dodelijke gevolgen altijd aanwezig zijn. Laboratoriumwerkers, in de academische wereld en de industrie, moeten een geschikt veiligheidsbewustzijn hebben om het risico tot een acceptabel niveau te brengen. Uit onderzoeksrapporten en reviews van laboratoriumongevallen blijkt echter dat er door een gebrek aan veiligheidsbewustzijn en risicogedrag nog steeds ongelukken plaatsvinden. Onvoldoende veiligheidstraining wordt vaak als één van de oorzaken van deze ongevallen gezien.

Vaak worden veiligheidstrainingen gegeven door middel van conventionele trainingsmethoden, zoals klassikale presentaties, video's en gedrukte handleidingen. Bij deze methoden moeten de leerlingen echter naar de instructeur luisteren of lange tekstuele paragrafen lezen. Ze zijn passief in hun leerproces en kunnen snel hun aandacht verliezen als ze niet gemotiveerd zijn, waardoor de veiligheidstraining niet effectief is. Virtual reality (VR) technologie kan worden gebruikt als een praktisch hulpmiddel om gevaarlijke situaties te simuleren in een virtuele omgeving waar leerlingen hun veiligheidsbewustzijn kunnen trainen zonder zichzelf of anderen echt in gevaar te brengen. Een combinatie van deze technologie met een 'serious game'-aanpak zou veiligheidstraining motiverender en boeiender kunnen maken dan conventionele trainingsmethoden.

Daarom gaat dit doctoraatsonderzoek de immersieve virtual reality technologie exploiteren om leerervaringen in een virtuele laboratorium omgeving uit te voeren, waar laboratoriumveiligheid kan aangeleerd worden zonder mensen echt in gevaar te brengen. In dit geval, dit onderzoek zou een beter begrip willen verkrijgen hoe werknemers gemotiveerd en geengageerd zijn om veiligheidstrainingen te volgen in een immersieve VR omgeving. Het ontwerpen en ontwikkelen van een VR serious game is echter zeer uitdagend. Er zijn veel factoren waarmee rekening moet gehouden worden voor een optimaal trainingsspel. Daarom is het doel van dit onderzoek niet enkel om te verifiëren dat dit immersieve hulpmiddel veiligheidstrainingen motiverender

en boeiender kunnen maken, maar ook om richtlijnen op te stellen hoe men een VR serious game kan ontwerpen, ontwikkelen en implementeren in een veiligheidstraining. In dit proefschrift, is er een voorbeeld getoond van het ontwerp- en ontwikkelingsproces van een VR serious game, genaamd VR LaboSafe Game. De opgestelde richtlijnen zijn generiek en kunnen toegepast worden op het ontwikkelen van andere gelijkaardige VR veiligheidstrainingen. Deze richtlijnen bevatten de integratie van instructieontwerp, gameontwerp en VR-ontwerp. Ook worden online bronnen en optimalisatietechnieken voorgesteld om zelfstandig een VR-game te kunnen ontwikkelen.

Tijdens de ontwikkeling van VR LaboSafe Game, werd het spel getest op academische en industriële populatie om de gebruiksvriendelijkheid van het systeem en de simulatorziekte te analyseren. Het is zeer belangrijk dat het spel gemakkelijk te gebruiken is en weinig simulatorziekte veroorzaakt vooraleer er verdere evaluatietesten worden gedaan. De eerste test, dat werd uitgevoerd met de eerste versie van VR LaboSafe Game 1.0, toonde aan dat wanneer instructies enkel via tekst worden vertoond, het VR spel eerder een slecht gebruiksvriendelijkheid heeft. Deze test toonde ook aan dat langdurige, continue ervaringen in VR (>40 minuten) zal leiden tot milde symptomen van simulatorziekte. Daarom, is het VR spel verder aangepast om tekstuele informatie te verminderen, gesproken instructies toe te voegen en om begeleidende animaties te vertonen. Spelers zijn ook toegelaten om pauzes te nemen van VR tussen de spelniveaus. Een test werd uitgevoerd met de geüpdatete VR LaboSafe Game en toonde inderdaad aan dat de gebruiksvriendelijkheid en simulatorziekte verbeterd waren vergeleken met versie 1.0 van het spel. Deze VR LaboSafe Game 2.0 is dan geschikt om te gebruiken in verdere studies.

In een evaluatieonderzoek met VR LaboSafe Game werden eerder persoonsgerichte variabelen onderzocht om te bepalen hoe gemotiveerd en geëngageerd medewerkers van een chemisch bedrijf zijn voor veiligheidstrainingen met een conventionele methode en met een VR serious game. Vanuit dit experiment blijkt dat laboranten en managers eerder persoonlijk gemotiveerd zijn om veiligheidstrainingen te volgen voor hun eigen veiligheid en die van anderen. Ze vinden veiligheidstraining echter repetitief, passief en gemakkelijk om de aandacht te verliezen wanneer deze wordt aangeboden met conventionele trainingsmethoden. De resultaten tonen aan dat werknemers meer intrinsiek gemotiveerd zijn om veiligheidstraining te volgen met een VR serious game dan met conventionele trainingsmethoden. VR is voor sommigen echter nog te ingewikkeld en onbekend waardoor dit de autonome motivatie kan beïnvloeden. Vooral oudere werknemers (ouder dan 50 jaar) hebben meer moeite met het gebruik van VR-headsets dan jongere werknemers (jonger dan 30 jaar). Ondanks deze complicatie willen medewerkers meer VR-veiligheidstrainingen volgen dan met een conventionele methode. Er wordt voorgesteld om conventionele methoden te combineren met VR als een aanvullend hulpmiddel en om frequentere en kleinere sessies aan te bieden, waarbij geleidelijk VR-technologie wordt geïntroduceerd voor beginners.

De resultaten van dit proefschrift zijn tweeledig: 1) er wordt een reeks van richtlijnen gedemonstreerd over het ontwerp, de ontwikkeling en implementatie van VR serious games voor veiligheidstrainingen, en 2) er wordt een functioneel artefact gecreëerd om veiligheidsbewustzijn in laboratoria aan te leren. Met dit artefact is het bewezen dat een VR serious game de intrinsieke motivatie en engagement van de leerling kan verbeteren. Echter, vooraleer men VR trainings implementeert, zou men eerst rekening moeten houden om mensen gewoon te laten worden met de VR technologie.

List of abbreviations

Abbreviation	Description
6DOF	6 degrees of freedom
ADDIE	Analyze, design, develop, implement and evaluate
AI	Artificial intelligence
AR	Augmented reality
CAVE	Cave automatic virtual environment
CRRA	Centre de recherche Rhône-Alpes
DBR	Design-based research
ESR	Early stage researcher
ETN-CHARMING	European training network for chemical engineering immersive learning
GUI	Graphical user interface
HMD	Head-mounted display
IDE	Integrated development environments
NPC	Non-player character
NUI	Natural user interface
PPE	Personal protective equipment
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
RAMP	Recognize hazards, assess risks, minimize risks, prepare for emergencies
SDS	Safety data sheet

Abbreviation	Description
SDT	Self-determination theory
SSQ	Simulator sickness questionnaire
SUS	System usability scale
VR	Virtual reality
WP	Work package

Contents

CHAPTER 1

General Introduction

1.1 Introduction

Training interventions on the topic of health and safety are of major importance for any workplace where hazardous chemicals are involved. This holds especially true for chemical plants and chemical laboratories. Several reviews and investigations of past accidents reveal that insufficient safety training was one of the causes for accidents both in chemical process industry (Bhusari et al., 2020; Dakkoune et al., 2018) and in academic research laboratories (Chen et al., 2020; Gopalaswami & Han, 2020; Simmons et al., 2017). Laboratory personnel have to work with a broad variety of hazardous substances and many different kinds of equipment, which could lead to a high risk of injury or even death, if these are not well managed (Schröder et al., 2016). Unfortunately, it has been noted that lack of safety awareness and at-risk safety behaviour are still prevalent in research laboratories (Chen et al., 2020; Gopalaswami & Han, 2020; Papadopoli et al., 2020; Walters et al., 2017). Therefore, adequate safety training is a necessity and is an inherent part of any safety management system. In a recent review article in the journal Nature Chemistry, Ménard and Trant (2019) argued that academic lab safety is still underdeveloped and urged a call for action to perform more safety-related studies in academic laboratories, including studies to evaluate and improve the effectiveness of current safety training.

Current safety training interventions are commonly provided using traditional teaching methods, such as classroom lectures, videos and printed safety manuals. These methods are easy to develop, are able to teach a large amount of information to a large number of trainees at once and offer great group control (Withers et al., 2012). However, these methods include a unidirectional flow of information where the trainee is required to pay attention and listen to the instructor (Bhide et al., 2015). As such, a low engagement of the trainee can lead to boredom and diminished attention to the learning content, thus making the training less effective (Fivizzani, 2005). According to a marketing survey of 150 health and safety professionals in 2003, boredom is the main obstacle of effective safety training (Gronbacher, 2004). Several contributing factors that could invoke this boredom are: 1) program topic repetitiveness; 2) low perception of importance due to overconfidence; 3) non-engaging presentations or videos; 4) and too much focus on theoretical safety knowledge and compliance to safety rules (Chen et al., 2020; Elston & Luttrell, 1997; Gronbacher, 2004). In general, researchers and safety experts agree that there is a need for safety training programs utilising more engaging learning methods and incorporating competency-based skills development (Alaimo et al., 2010; Elston & Luttrell, 1997; Laberge et al., 2014). Other common safety training methods are on-the-job and hands-on training, where the trainee learns the necessary safety measures by hands-on activities supervised by more experienced workers. This method encourages the trainees to be active in their learning process and cultivates their decision-making skills through experience (Bhide et al., 2015). However, training of highly dangerous situations is not allowed with this method because this puts them and others at a high risk.

In the last decades, the rapid growth of technological innovations has enabled the development of computer-based technologies that can provide realistically simulated experiences. These simulation-based technologies, such as immersive virtual reality (VR) and digital games, have been creating great opportunities to improve safety training methods. With these technologies, the trainee can be trained in a realistic representation of the workplace environment, performing realistic tasks with a high degree of interaction (Checa & Bustillo, 2019). This makes it possible to improve decision-making skills on important safety issues, where mistakes can be made without real-life hazardous consequences. The activity of learning-by-doing and learning from mistakes can make the learning experience more engaging and more memorable (Miliszewska & Sztendur, 2011). Also, practical procedures can be virtually rehearsed without the need of a second observer and without the high cost related to the materials and equipment. Simulations and VR technologies are being used successfully in various safety domains, such as mining (Grabowski & Jankowski, 2015), construction (Li et al., 2018; Sacks et al., 2013), aviation (Buttussi & Chittaro, 2018), and fire safety (Smith & Ericson, 2009). However, studies of VR-based solutions in the field of chemical safety are still limited. Some studies have reported using virtual reality for the training of chemical plant operators (Garcia Fracaro et al., 2022; Patle et al., 2018) and laboratory safety (Makransky, Borre-Gude, & Mayer, 2019; Srinivasan et al., 2022). In addition, there is still a lack of studies evaluating the effectiveness of using VR in chemical safety training (Garcia Fracaro et al., 2022; Srinivasan et al., 2022). More particular, there is a need for more learner-centred research of training activities (Burke et al., 2006; Casey et al., 2021). The term *learner-centred* (or *person-centred*) teaching methods primarily focus on the psychological needs of an individual learner, rather than on the training content, technology and assessment (Kebritchi & Hirumi, 2008). These learner-centred methods incorporate the understanding of how the human mind works and how people are motivated and actively engaged in learning activities (Mayer, 2014a). A serious game design, for example, could make VR training more engaging and interactive for the learners. In addition, little information is found on the complicated process of designing and developing effective VR applications for safety training. Therefore, the current doctoral research satisfies this need by describing the process of design and development of such VR application and by evaluating the training motivation and engagement of using the developed VR serious game for chemical lab safety training.

1.2 Objectives and scope of research

This doctoral research project exploits immersive virtual reality technology to fully immerse the learning in a virtual laboratory environment, where laboratory safety is trained without putting the learner in danger. In particular, this research wants to obtain a better understanding of the motivation and engagement of employees following safety training in an immersive VR environment. Serious game is used to create this learner-centred training environment. The main objectives are: (i) to formulate design guidelines for VR serious games on lab safety training, (ii) to implement these in the development of such game, and (iii) to verify if increased motivation and engagement can be expected with this immersive training method. Moreover, This dissertation could also serve as a guide for the process of design and development of such immersive tools. Although an example is shown of a chemical lab safety training, the process of creating and implementing VR serious games can be generalised for other safety training programmes. Furthermore, one of the outcomes of this doctoral research is a fully operational VR prototype for chemical laboratory safety training.

1.3 ETN-CHARMING project

This doctoral research project frames within the ETN-CHARMING project, which stands for European Training Network for Chemical Engineering Immersive Learning (website: charming-etn.eu). This project takes on the challenge to develop learning strategies, content and prototypes for the application of games and virtual/augmented reality to motivate, teach and train children, students and employees in chemistry, chemical engineering and chemical operations. This inter-sectorial and interdisciplinary research project consisted of leading universities and industry participants and trained 15 Early Stage Researchers (ESRs) in the areas of innovative chemical engineering, instructional psychology & pedagogy and immersive technology. In particular, the current doctoral research is part of Work Package 3 (WP3), which focused its research on utilising VR technologies to train employees of the chemical industry on the topics of process plant operations and chemical laboratory safety. A close collaboration was established with other ESRs from the WP3 team (See Figure 1.1). Two VR applications were developed and evaluated: a VR training game for laboratory safety (VR LaboSafe Game) and a VR training for chemical operators (Operate Your Own Reactor). The current doctoral research investigates the motivation and engagement of VR training using the VR LaboSafe Game. Sofia Garcia Fracaro (Merck) investigated the learning effectiveness of VR training and Yusra Tehreem (Hochschule Emden/Leer) investigated the effect of sitting and standing VR experiences using the Operate Your Own Reactor. Ryo Toyoda (Newcastle University) established the in-game assessment framework and Tim Gallagher (Utrecht University) evaluated the learning analytics representations of both VR applications.

Most of the work in this dissertation was coordinated and performed in the Centre de Recherche Rhône-Alpes (CRRA) of the company Arkema. Arkema is a multi-national chemical enterprise based in France with leadership positions in specialty materials. The activities of researching and manufacturing these specialty materials are divided in four main business segments: Adhesive Solutions (construction & consumer, industrial assembly), Advanced Materials (high performance polymers, performance additives), Coating Solutions (coating resins, coating additives) and Intermediates (fluorogases) (Arkema, 2022b). Employees of Arkema have to work with a high variety of advanced and hazardous chemicals, such as hydrogen fluoride, acrylonitrile, organic peroxides, and many more. This means that they are exposed to a wide range of different safety hazards: from toxic to corrosive, and even explosive hazards. Working with such safety risks, requires Arkema to have a strong focus on health and safety of its employees and contractors. A core aspect of Arkema's safety process is the development of a common safety culture that raises the awareness of every employee (Arkema, 2022a). Rigorous safety training programmes are provided to develop and implement this safety culture. For this reason, VR technology can help employees of Arkema to train and get acquainted with handling specific safety hazards of the dangerous chemicals that they need to work with on a daily basis. Further, a serious game approach can sustain their engagement during the safety training programmes.

In order to create the VR serious game in this doctoral research, expertise was acquired from four partner universities of the CHARMING project: IT University of Copenhagen (ITU), Bielefeld University (UBI), Hochschule Emden/Leer (EMD), Utrecht University (UU). In a month-long period per university, called a secondment, intensive personal coaching was received at the university site or online in order to transfer interdisciplinary skills and knowledge about game design, instructional design and game development. Prof. Daniel Cermak-Sassenrath from ITU with ESRs Silvia Fornós and Nina Croitoru provided essential advice on game design principles. Prof. Liesbeth Kester and Prof. Lisette Hornstra from UU with ESRs Michaela Arztmann, Yuanyuan Hu and Tim Gallagher, provided expertise on instructional design principles and motivational psychology. Prof. Thies Pfeiffer from UBI/EMD with ESR Yusra Tehreem, organised workshops to obtain game development skills needed to create the VR LaboSafe Game.

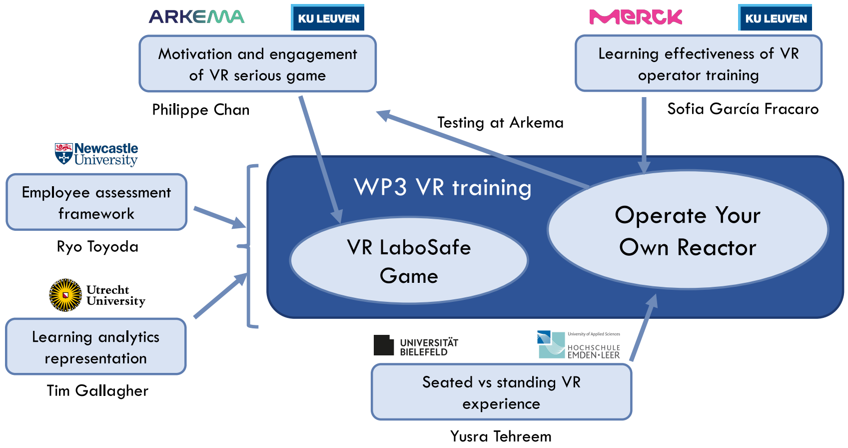


Figure 1.1: A representation of the international and interdisciplinary collaboration of WP3 ESRs: Philippe Chan (Arkema & KU Leuven), Sofia Garcia Fracaro (Merck & KU Leuven), Tim Gallagher (Utrecht University), Ryo Toyoda (Newcastle University) and Yusra Tehreem (Hochschule Emden/Leer & Bielefeld University)

1.4 Research methodology

The progress of this doctoral research study follows the design-based research (DBR) methodology. DBR is a type of methodology used in learning sciences, more specifically with educational technologies. Diverse DBR methods exist, but this thesis is grafted on the work Van Wyk and Villiers (2014) that presents a synthesised approach and exemplified it to a VR safety training for the African mining industry. This research methodology is systematic, iterative and is focused on improving educational practices. It aims to make both practical and scientific contributions by designing and developing solutions to solve complex real-world problems. For this reason, this methodology often leads to dual outcomes: on one hand, theoretical frameworks or design guidelines are established, and on the other hand, practical artefacts are created as a usable solution for real problems. DBR has been extensively used for studies involving the development of e-learning and e-training (Villiers, 2012). Hence, the DBR methodology is particularly suitable for this doctoral research study, because in this study, a VR application is produced for the purpose of chemical lab safety training, while also guidelines are determined for the process of the design, development and implementation of such artefact.

The process cycle of DBR is heavily based on the well-known instructional design model, called ADDIE. The ADDIE process consists of five phases to design instructional interventions: Analyze, Design, Develop, Implement and Evaluate (Branch, 2009). This model has some similarities with the process of commercial game development and is also used in the context of serious game development (Braad et al., 2016). The difference between these similar process models is that ADDIE is particularly used in the context of instructional design, while game cycle processes are used in the context of game development. In the design and development of serious games, ADDIE can be used to guide the process of creating an instructional method, while also the process of developing a game software is taken into account (Dimitriadou et al., 2020). For this reason, the DBR process cycle is a suitable fit for this research as it has a similar process plan but more focused on research outcomes. The DBR methodology makes use of these phases and applies them in a research context with distinct outcomes for each phase (van Wyk & de Villiers, 2014):

- ➢ *Problem analysis within context:* the state of the art and literature are reviewed in order to identify and analyse the real-world problem. The outcome of this phase is a research proposal with clear research goals;
- ➢ *Design solution:* the design process is initiated in order to solve the identified problem. Multiple perspectives are required in the design in order to tackle the complex nature of real-world situations. The outcome

of this phase is a well-defined design of the solution that is ready to be developed;

- *Develop solution:* the development process is initiated to develop the design of the solution into a practical artefact. Technological tools and technical skills are required for the development. This process consists of iterative cycles of testing and systematic refinement. The outcome of this phase is an innovative, operational prototype that is ready to be evaluated;

- *Evaluate in practice:* the developed artefact is tested and evaluated in a real-world context. Different research methods are used to collect data, this data is then carefully analysed in order to form conclusions;

- *Reflection:* the whole process of designing, developing and evaluating the prototype is reviewed, resulting in dual outcomes. In the context of practical contribution, the prototype can be implemented to solve real-world problems, which can be further refined and improved by initiating another DBR cycle. In the context of theoretical contribution, guidelines and design principles are established based on the experiences achieved during the entire process. Also these design principles can be further refined in multiple DBR cycles.

1.5 Thesis structure

A schematic overview of the chapters in this dissertation is presented in Figure 1.2. The structure of this dissertation follows the chronological process of the design-based research process according to Van Wyk and De Villiers (2014).

In **Chapter 2**, a systematic literature review is presented on virtual chemical laboratories. The state of the art is investigated concerning the research, technologies and instructional designs of these virtual environments. This chapter is related to the Problem analysis phase.

The Design and Develop phases of DBR are combined and described in **Chapter 3**. It contains the design justifications and a description of the development process for a VR serious game for safety training in chemical laboratories, called VR LaboSafe Game.

As part of the Develop phase, **Chapter 4** describes the implementation and tests of the VR LaboSafe Game on academic and industrial populations. These tests investigate the system usability and simulator sickness to ensure that the VR LaboSafe Game is an easy-to-use application that does not make users nauseous while using the VR headset. Specific design decisions of virtual environments are crucial to control the usability and learnability of the VR training game. Therefore, based on the results of the usability testing and participant observations, the VR game is adjusted and refined in iterative cycles in order to optimise the usability and simulator sickness. A version of the training game is achieved with satisfactory results so that this version is suitable for further studies.

In **Chapter 5**, the VR LaboSafe Game is evaluated on employees of Arkema to determine its impact on motivation and engagement compared to more conventional training methods. In the remainder of this dissertation, 'conventional training method' is defined as classroom and lecture based methods without extra engaging elements (e.g. interactive instructor, digital tools). Particularly in this chapter, the compared conventional method is a video lecture. Using questionnaires and interviews, person-centred insights are gathered about the use of VR serious games for training purposes. This chapter is related to the Evaluate phase of DBR.

Finally, **Chapter 6** reviews the previous chapters and forms a general conclusion about the entire progress of this doctoral research. A list of guidelines is presented to guide future practitioners on the design, development and implementation of VR serious games for safety training. Future perspectives on research and exploitation are also proposed in this chapter.

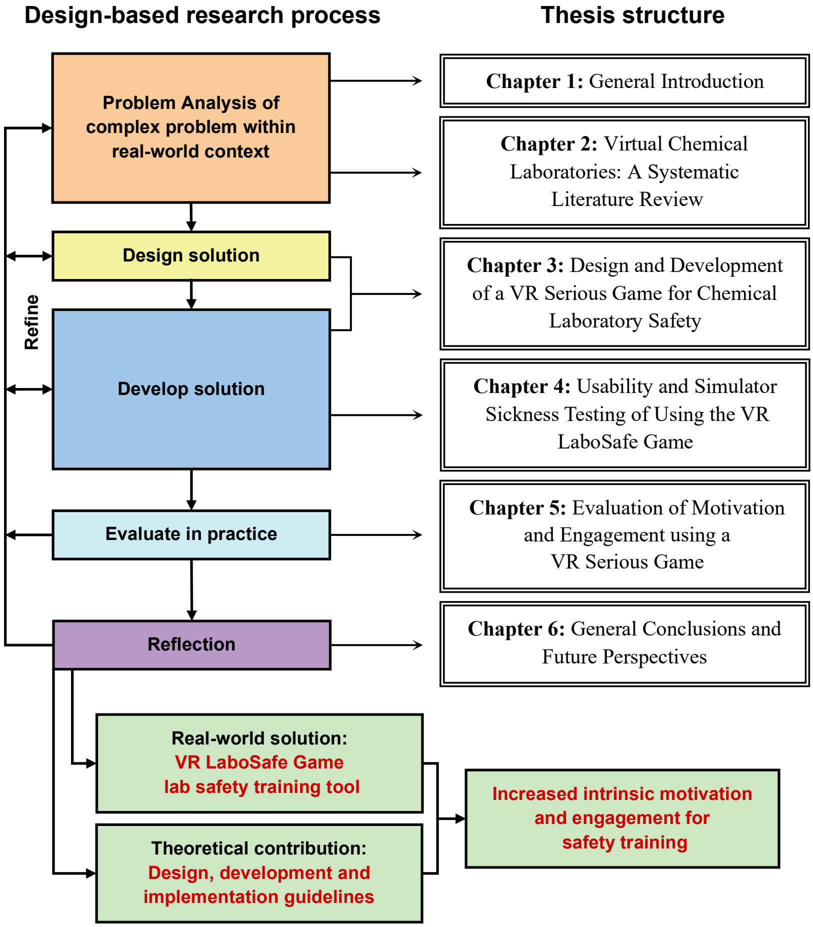


Figure 1.2: Graphical overview of this dissertation based on the process of design-based research described by (van Wyk & de Villiers, 2014).

CHAPTER 2

Virtual Chemical Laboratories: A Systematic Literature Review

In the first stage of the design-based research process, an initial analysis is performed in order to determine the main problem that the research should be focusing upon. For this reason, this chapter reviews the state of the art of research literature about virtual chemical laboratories. In order to thoroughly investigate the literature landscape, a rigorous systematic literature review procedure is applied. However, because of the scarcity of literature work on virtual labs specifically for safety training, the scope of the literature review is expanded to include virtual labs for education of laboratory practices. Notwithstanding the small number of papers, a clear insight is given in this chapter on the research, instructional designs and technologies of existing chemical virtual laboratories.

This chapter is published as:
Chan P., Van Gerven T., Dubois J.-L., Bernaerts K. (2021). Virtual chemical laboratories: a systematic literature review of research, technologies and instructional design. Computers & Education Open, Volume 2.
Author's contribution: Chan P. performed the literature search and coding, and drafted the manuscript. Bernaerts K., Dubois J.-L. and Van Gerven T. managed and supervised the project.

2.1 Introduction

Laboratory work is often seen as an essential part of chemistry education. Reid and Shah (2007) presented four important skills that students acquire during practical laboratory sessions: (1) skills related to learning chemistry, (2) practical skills, (3) scientific skills, and (4) general skills. Seery (2020) further elaborated that laboratory work is distinct from the rest of the curriculum in a way that a laboratory is “a complex learning environment, whereby students need to draw together constituent skills, including learning the requisite practical skills, and knowledge, and applying them to a scientific task”. He stated that the laboratory is “the place to learn how to do chemistry”. However, physical laboratory sessions are labour- and time-intensive for the personnel involved and the laboratory infrastructure is highly expensive (Bretz, 2019; Reid & Shah, 2007). So making these practical sessions available is sometimes a challenging task, especially during a pandemic lockdown when these facilities are not accessible. Online digital tools, such as videoconferencing applications, e-learning platforms and online videos, have been used as alternative to teach the chemical theory behind lab experiments, but a major challenge still exists to adapt practical exercises.

Virtual laboratories are one of the digital tools that can be used to provide distance learning for laboratory sessions. These virtual labs are computer-simulated learning environments that can range from simple 2D visualisations of laboratory experiments to advanced 3D simulations that try to replicate real laboratory environments (Jones, 2018). With recent virtual reality technology, it is even possible to be fully immersed in the virtual laboratory environment performing realistic laboratory handling (Han et al., 2017; H. Kim et al., 2019). Some benefits that virtual laboratories can offer, compared to traditional hands-on laboratories are (Ali & Ullah, 2020; Alkhaldi et al., 2016; Faulconer & Gruss, 2018): reduced cost, greater accessibility, time-saving, safe environments, and flexibility of self-regulated learning. However, depending on how the virtual lab is used, absence of other students or tutors, and lack of real-life feel of a laboratory may present drawbacks of these virtual applications (Lynch & Ghergulescu, 2017).

Due to distance learning becoming more popular nowadays, one can expect more varieties of virtual laboratories in the near-future. However, designing and developing such a complex virtual learning environment is not always that easy. It often requires a multidisciplinary team with different levels of expertise (e.g. computer scientists, educational technologists and chemistry teachers) in order to create an effective learning experience (Mikropoulos & Natsis, 2011). Furthermore, research has shown that the technological aspect is not the only contributing factor for the design of effective virtual learning environments.

In some cases, the technological design can even be inhibiting for cognitive learning processes if not optimally designed (Makransky, Terkildsen, & Mayer, 2019; Mayer, 2014a). A rigorous instructional design is required that utilises well-established learning theories and instructional support in order to optimise the effectiveness of the virtual laboratory experience.

The main goal of this systematic literature review is to provide an extensive overview of previous research on virtual laboratories in chemistry education. We investigate three main characteristics: research methods, technology and instructional design. Therefore, three research questions (RQs) are stated for this study:

- RQ 1: What are the main research purposes, evaluation methods and learning outcomes in studies on using virtual chemistry laboratories for educational purposes?
- RQ 2: Which technologies have been used for virtual chemical laboratories and what is the current trend?
- RQ 3: What learning theories and instructional design features have been applied in virtual chemical laboratories?

Eventually, this review could contribute as an aid for teachers and educational developers to select effective solutions for the distance learning of chemical laboratory practices.

2.2 Related work

Research on virtual laboratories is not a new topic. In fact, several reviews have been published comparing virtual and remote laboratories to traditional hands-on laboratories (Brinson, 2015; Faulconer & Gruss, 2018; Lynch & Ghergulescu, 2017; Ma & Nickerson, 2006). However, these reviews include laboratory practices of many other disciplines (e.g. biology, physics, engineering sciences) and few virtual laboratories in chemistry are mentioned. Only four other reviews were found that discuss virtual chemical laboratories more in depth (Ali & Ullah, 2020; Bellou et al., 2018; Sypsas & Kalles, 2018; Tatli & Ayas, 2010).

Tatli and Ayas (2010) published the first review on the subject and examined 13 papers reporting virtual chemical laboratories that are based on a constructivist learning approach in order to analyse their advantages and disadvantages. They investigated the purpose of the studies, sample size, data collection tools and study results. This study concluded that these labs allow students to focus on the process rather than on the equipment, promote active participation with little to no time waste, and allow experiments to be repeated in a safe environment. The major drawback was that students using virtual laboratories could not feel, smell or touch as in a real laboratory.

Sypsas and Kalles (2018) analysed 29 peer review articles of virtual laboratories in the domains of biology, biotechnology and chemistry. They were focused on its effectiveness as supplementary tool and the educational approaches that are used. This study concluded that virtual laboratories show similar or better results than conventional methods for secondary education and that they are most effective when it is combined with real laboratory for post-secondary education. Blended learning and inquiry learning were the most used educational approaches. They also mention that their review might have excluded a lot of research papers, and they encourage to update such reviews frequently as technological advancements improve rapidly.

Bellou et al. (2018) reviewed 43 studies of digital learning technologies in primary and secondary chemistry education. They primarily looked into the learning technologies, pedagogical approaches, research methods and learning outcomes of the studies. From the technical approaches they have analysed, seven of them were virtual labs, whereas the most used technological approaches were multimedia and simulations. Their findings of their review suggest that most studies involve secondary education; cover mostly particulate nature of matter as topic; and have adopted mostly constructivist learning theories. Furthermore, the research method of most studies assessed the student's knowledge following an experimental and quasi-experimental design with a majority reporting positive learning outcomes. However, the authors have remarked that proceedings of conferences and book chapters were not included and suggested a more systematic effort with more meta-analyses of empirical studies.

Ali and Ullah (2020) conducted a literature review collecting 42 different virtual chemistry laboratories. They proposed a classification of the type of graphical interfaces used in these virtual laboratories. The authors made a distinction between 2D, 3D and video metaphor virtual chemistry laboratory with a further separation of offline and online virtual labs. From this collection, a comparison was made between 2D and 3D virtual chemistry laboratories to reveal their similarities and differences. Ali and Ullah (2020) noted that 2D virtual labs lack realism and provide low immersion compared with labs that use 3D graphical interface. Furthermore, they have discovered that most of the virtual labs did not provide any guidance on the procedure of an experiment.

The work in this chapter is distinct from these literature reviews. Firstly, because this study includes immersive virtual reality technologies. Ali and Ullah (2020) briefly mentioned this type of technology, but it was classified under 3D virtual laboratories. Secondly, this study provides a holistic overview of previously used virtual chemical laboratories in research literature. This means that it also includes conference proceedings and virtual laboratories at university level, which were not included in the studies of Sypsas et al. (2018) and Bellou et al. (2018), respectively. Thirdly, this study considers features of instructional design that were taken as the basis for the design of the virtual chemical laboratories.

2.3 Methodology

In order to conduct this systematic literature review, we have followed PRISMA's principles and guidelines (Moher et al., 2009). These guidelines help researchers to conduct a transparent and complete reporting of systemic literature reviews. It requires the author to specify the search strategy, eligibility criteria, selection process and data collection process.

2.3.1 Database and search keywords

The first step of this systematic literature review is the literature search in an online the database. As such, a search was initiated in November 2020 using Web of Science as scientific database. A combination of search terms was used to find publications about virtual applications or games that considered chemical laboratory, chemical experiment or laboratory safety instructions:

- (virtual OR game) AND (chemical OR chemistry) AND (laboratory OR lab);
- (virtual OR game) AND (“chemical experiment” OR “chemistry experiment”);
- (virtual OR game) AND (“lab safety” OR “laboratory safety”).

These search terms should be found in the title, abstract or list of keywords of publications between the years 2000 to 2020. This search yielded 806 records after removing duplicates. Additionally, 8 records were added by manually searching on Google Scholar and by examining the references of the publications. Eventually, a total of 814 records remained to be screened.

2.3.2 Inclusion and exclusion criteria

The next step of our study is dedicated to selecting the relevant articles to be included for this review by screening the title and abstract of each record. For this selection, a number of inclusion and exclusion criteria were used in order to filter out the irrelevant publications (See Table 2.1). This title and abstract screening resulted in 113 valid publications and 701 records were removed. These selected publications were then subjected to further screening of their full text. Publications were filtered out of virtual laboratories that did not contain a chemical experiment or a representation of a virtual laboratory environment. For example, publications where a virtual application was used to only visualise molecule structures (Ferrell et al., 2019) were excluded. Another criterium is that the type of display technology (e.g. 2D, 3D or immersive VR) should be described in the text or represented in the images of the publication. After screening of these texts, a total of 76 publications remained that are included in this literature review.

Table 2.1: A list of inclusion and exclusion criteria used to select relevant articles from the search results.

Inclusion	Exclusion
Journals and conference proceedings	Reviews, abstracts and non-peer reviewed publications
Virtual laboratories used for chemistry education	Publications with full text that are not accessible
Contains chemical laboratory practices or laboratory safety	Virtual applications that are only used to teach chemical concepts (e.g. molecule visualisation, periodic table)
Uses 2D, 3D graphical interfaces or immersive virtual reality devices	Virtual lab applications that require the real environment (e.g. augmented reality)
Publications must be in English	Publications that are not in English

2.3.3 Data analysis and coding

In this last step, the 76 included publications were analysed by coding the relevant information that is appropriate to our research questions onto a spreadsheet. These variables were then classified in distinct categories introduced in this section. Appendix A.1 shows a table of this classification with clear description of each category. The full coding scheme of each publication with their identification number can be found in Appendix A.2.

Research purpose

The research purpose of the publications that report the use of virtual chemical laboratories can be classified into three main categories: comparative, evaluative and technical.

Comparative studies investigate two or more intervention groups either comparing the media or the design of the virtual laboratory. Following the definitions of Mayer et al. (2014a), the former is called media comparison while the latter is called value-added research. In media comparison research, the learning outcome of an experimental group using a virtual chemical laboratory application is compared with a control group that had the same learning content but with a different educational medium. For example, the study of Tarng et al. (2017) compared the experimental group performing experiments in the virtual laboratory with a control group performing experiments in a real hands-on laboratory. Furthermore, lectures, videos, text and demonstrations were grouped together as 'passive media' because these media do not require active participation from the participant, unlike virtual and hands-on laboratories. In value-added studies, a basic version of a virtual laboratory application is tested with a control group, while the intervention group uses the same basic version but with one design feature added or changed. For example, the study of Ullah et al. (2016) used a virtual lab with procedural guidance and a virtual lab without procedural guidance. The learning outcomes of both versions are then compared to each other which allows the investigation of the effectiveness of a specific design principle.

Evaluative studies only consider the virtual laboratory group in order to evaluate a particular outcome. Affective reactions, such as attitude, satisfaction or self-efficacy of the participant can be measured in order to evaluate the user experience and usability of the system. In this case, we have grouped these publications under 'user study'. Other studies have investigated the performance of using the virtual laboratory without a control group only to evaluate the performance gain or assessment method. This category is called 'performance assessment'. Evaluative studies with other purposes, for example a correlation study (Scherer & Tiemann, 2012), are identified with 'Other'.

Technical studies do not perform any measurements to evaluate the virtual laboratory, but rather describe its design and development. Although no measurable results are presented in these publications, they are still valuable for this review as they describe the technical advances that has been implemented in virtual chemical laboratories.

Evaluation methods

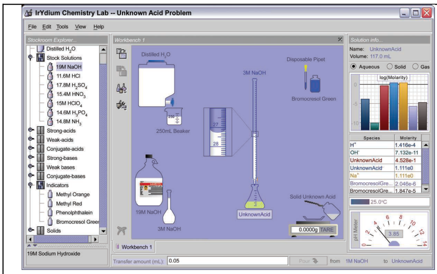
Evaluation methods are collected and categorised from studies who have performed measurements (i.e. comparative and evaluative studies). These include quantitative and qualitative methods, similar to Brinson et al. (2015): test, lab practical, real-time assessment, school grade, questionnaire, interview and observation. *Lab practical* refers to hands-on lab experiments to assess laboratory skills, while *real-time assessment* refers to collecting data within the virtual lab application that are retrievable using log files.

Technology type

A distinction is made between display technologies and natural user interfaces (NUIs). Display technologies show graphical representations using a certain display device and are categorised in 3 different types: 2D desktop, 3D desktop and immersive VR. Studies belonging to the *2D desktop* category describe virtual chemical laboratories that are displayed on a desktop monitor display and feature a 2D representation of the environment and objects. In the *3D desktop* category, virtual chemical laboratories are displayed on a desktop monitor display as well, but are 3D in nature, meaning that the virtual environment and objects have a depth and are built from 3D geometries. Studies included in the *immersive VR* category describe the use of modern VR devices where the user is fully immersed in the virtual environment without visual interaction with anything else from the real world other than the display. This mainly concerns the use of VR head-mounted displays (HMDs). These VR devices are also able to display a different image per eye, allowing a 3D stereoscopic view that results in the perception of real depth. Figure 2.1 shows examples of these types of technologies.

A further distinction can be made whether the authors have used special input devices as NUIs in addition to a display technology. These NUIs use human movement or gestures as input to control the system *"in such way that the user is not aware of the existence of an interface"* (Jagodzinski & Wolski, 2015). These include devices that provide advanced tracking capabilities, such as movement/rotational tracking, spatial tracking, and tracking of hand gestures or body gestures.

Figure 2.1: Examples of types of technology used in virtual chemical laboratories in the reviewed studies.



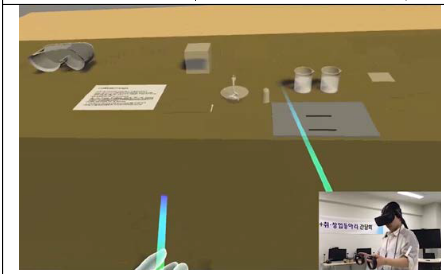
Name: UnknownAcid
Volume: 117.0 mL

Species	Molarity
H^+	1.416e-4
OH^-	7.132e-11
UnknownAcid	4.528e-1
$UnknownAcid^-$	1.111e0
Na^+	1.111e0
BromocresolGre...	2.046e-6
BromocresolGre...	1.847e-5

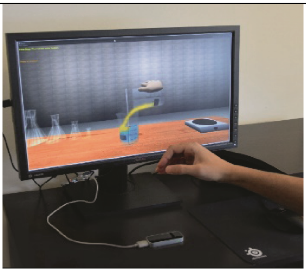
Transfer amount (mL): 0.05 from 1M NaOH to UnknownAcid

2D Desktop (Yaron et al., 2010)

3D Desktop (Su & Cheng, 2019)



Immersive Virtual Reality (Duan et al., 2020)



Natural User Interface (Aldosari & Marocco, 2016)

Instructional design

The instructional design of the virtual chemical laboratories is analysed by identifying the learning theory and instructional support elements that were used in these publications. This was done by examining which learning theory has been applied using the terms derived from the collection of Kebritchi and Hirumi (2008). Likewise, for examining instructional support elements, the list of Wouters and Oostendorp (2013) was used to identify the terms. Publication with no learning theory or no instructional support indicated, were marked with ‘not specified’.

2.4 Results

This section connects the results of the review inquiry to the three research questions mentioned above and is divided into three subsections: research methodology, technology and instructional design.

2.4.1 Research methodology

Research purpose

Table 2.2 shows the three categories of research purposes (i.e. comparative, evaluative and technical studies) with the corresponding publications as reference, while Figure 2.2 shows the relative distribution of these categories.

We have observed that the majority of the publications could be identified as ‘comparative study’ (n = 38, 50%). In this category, most studies have conducted media comparison (n = 33) rather than value-added research (n = 3), while some studies have combined both (n = 2).

The second most common research purpose is the ‘evaluative study’ category (n = 24, 32%). Most studies (n = 15) used this approach to conduct a user study by examining the affective reactions of the participants. Five other studies (n = 5) have investigated the performance of using the virtual laboratory without a control group in order to evaluate the performance gain or real-time assessment method. Some studies (n = 3) combined both user study and performance assessment. Also, one other study used their virtual chemical laboratory for a correlation study (Scherer & Tiemann, 2012).

Publications belonging to the ‘technical study’ category are found to be the least common among the two other categories (n = 14, 18%). These studies are intended to introduce the design and technology of the virtual chemical laboratory and to describe the development of such applications.

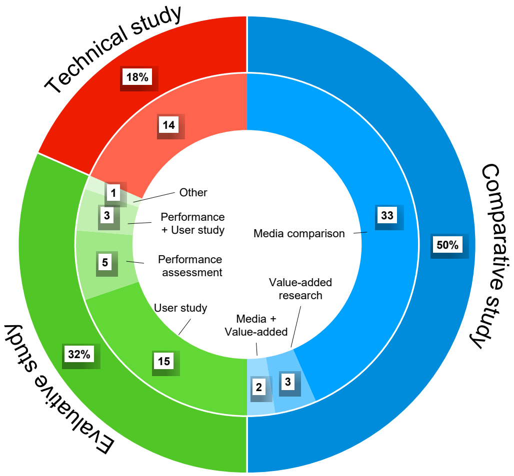


Figure 2.2: Pie chart presenting the distribution of publications according to their research purpose. Corresponding references are listed in Table 2.2.

Table 2.2: Classification of categories of research purposes.

Research purpose	Sub-category	n	References (ID no. from Appendix A.2)
Comparative study	Media comparison	35	1, 5, 10, 11, 14, 17, 18, 24, 25, 30-36, 38-40, 45, 47, 49, 53, 56, 57, 60-65, 67-70
	Value-added research	5	13, 39, 48, 65, 75
Evaluative study	User study	18	2, 4, 6, 7, 15, 19, 20, 22, 23, 41, 50, 52, 54, 55, 58, 71, 73, 76
	Performance assessment	8	8, 9, 16, 19, 26, 58, 72, 73
	Other	1	59
Technical study		14	3, 12, 21, 27-29, 37, 42-44, 46, 51, 66, 74

With n = number of studies

Evaluation methods

To evaluate the effectiveness of virtual chemical laboratories, a mix of quantitative and qualitative evaluation methods has been used for measuring cognitive, affective and/or skill-based learning outcomes. As seen from Figure 2.3 there is a notable difference in the use of these methods between comparative and evaluative studies.

Tests are the most frequently used evaluation methods for comparative studies (n = 32, 84%), while questionnaires are the most used methods for evaluative studies (n = 21, 88%). Comparative studies also seem to have used qualitative evaluation methods to measure affective outcomes of participants, such as questionnaires (n = 23, 61%), interviews (n = 8, 21%) and observations (n = 7, 18%). It is also observe that lab practicals were applied in comparative studies (n = 9, 24%) and not in evaluative studies, while real-time assessments were used more in evaluative studies (n = 6, 25%) than in comparative studies (n = 2, 5%).

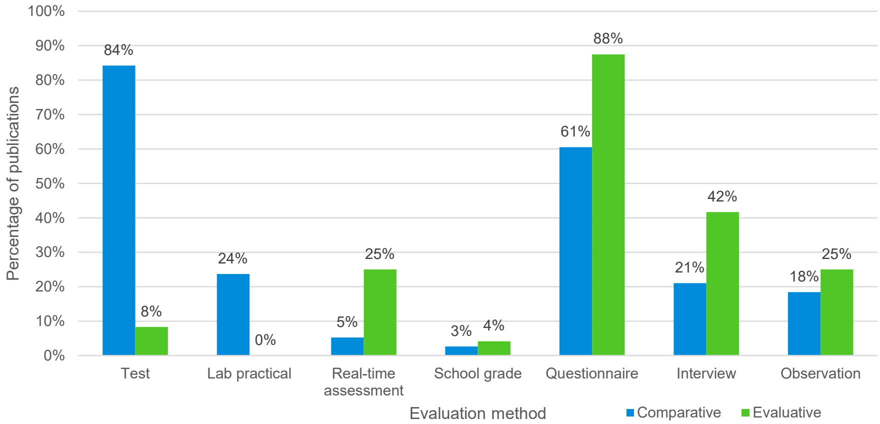


Figure 2.3: Bar graph showing the percentage of the evaluation methods used in comparative studies (blue) and evaluative studies (green).

Learning outcomes

When investigating learning outcomes, we observed that most comparative studies measure cognitive learning outcomes (n = 35, 92%), whereas most evaluative studies measure affective outcomes (n = 19, 79%), as seen in Figure 2.4. However, affective outcomes of participants are also frequently measured in comparative studies (n = 22, 58%). This is because these studies also evaluate the usability of the virtual laboratory and opinions of the participants besides cognitive and/or skill-based comparison. The least evaluated learning outcome is the skill-based outcome, both for comparative (n = 11, 29%) and evaluative studies (n = 2, 8%). When investigating the sub-categories of these learning outcomes in Table 2.3, declarative knowledge is found to be the most measured learning outcome in the cognitive domain (n = 34) and usability in the affective domain (n = 31). Studies with 'Other' category investigated items that are not outcomes of learning, such as level of constructivist teaching (Tatli & Ayas, 2012) and user profile (Annetta et al., 2014).

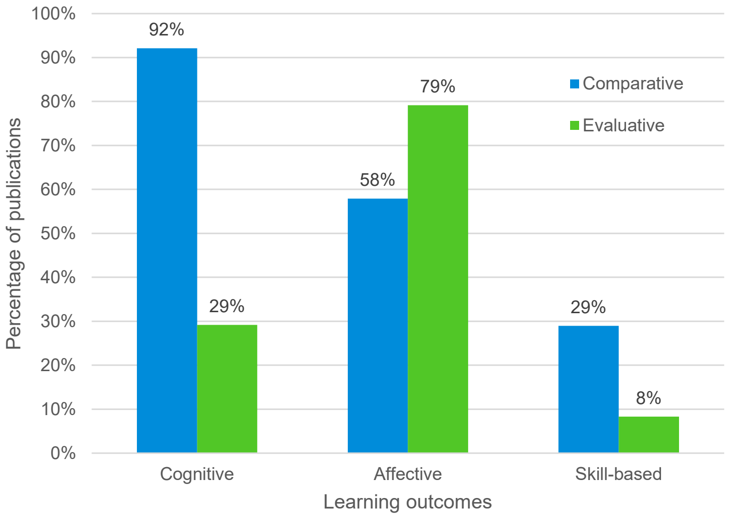


Figure 2.4: Bar graph of learning outcome measured in comparative studies (blue) and evaluative studies (green).

Table 2.3: Collection of learning outcomes covered in the reviewed publications.

Learning outcome domain	Sub-category	n	References (ID no. from Appendix A.2)
Cognitive domain	Declarative knowledge	34	1, 5, 8, 9, 13, 14, 17, 18, 24, 25, 30, 32-35, 38-40, 47-49, 53, 56, 57, 60-62, 64, 65, 67-70, 72, 75
	Procedural knowledge	11	8, 14, 16, 18, 25, 26, 49, 53, 58, 65
	Conditional knowledge	20	11, 13, 25, 26, 32-36, 38-40, 47-49, 56, 59, 67, 70, 72
Affective domain	Attitude	12	4, 10, 14, 17, 22, 25, 31, 32, 52, 53, 67, 69
	Usability	31	2, 5-7, 13, 15, 17, 19, 20, 23, 31, 39-41, 50, 52-55, 57, 58, 60, 62, 65, 67-69, 71-73, 76
	Self-efficacy	11	5, 17, 39, 40, 45, 47, 52, 53, 61, 65, 71
	Social presence	1	48
Skill-based domain	Practical laboratory skills	11	5, 19, 25, 30, 47, 49, 53, 57, 68, 69, 73
Other		2	9, 63

With n = number of studies

Media comparison

A popular discussion can be noticed within the reviewed studies about how virtual chemical laboratories are compared with traditional teaching methods. For this reason, the majority of the studies have conducted media comparison research. Table 2.4 shows an overview of the virtual labs (or combination with virtual labs) with the compared medium. Publications in which significantly better results are presented for using virtual labs are identified as ‘positive’, while publications showing non-significant different results are identified as ‘equal’. Publications showing significant results against virtual labs are identified as ‘negative’.

Results of these studies show that when compared to passive media, a majority of virtual labs report positive improvement, mainly in declarative and conditional knowledge (n = 11). Some studies (n = 4) reported equal effectiveness in declarative knowledge, but identified better results in conditional knowledge and self-efficacy in favour of virtual labs (Herga et al., 2015; Jagodzinski & Wolski, 2014; Makransky, Borre-Gude, & Mayer, 2019; Wolski & Jagodzinski, 2019). When virtual labs are compared to hands-on laboratories, results are mixed. Six studies (n = 6) reported positive improvement for virtual labs mainly in declarative knowledge with a total population of 639 participants, while nine studies (n = 9) observed equal effectiveness including declarative knowledge and skill-based learning outcomes with a total population of 1662 participants. Also for affective learning outcomes, virtual labs are not significantly different from hands-on labs in terms of attitude towards chemistry laboratory (including anxiety, satisfaction), usability and self-efficacy (regarding confidence) of the participants (Dalgarno & Lee, 2009; Hensen et al., 2020; Winkelmann et al., 2014). Two publications presented worse results in attitude and usability against the virtual lab compared to hands-on laboratory (Hensen & Barbera, 2019; Hensen et al., 2020) and one study reported worse results in declarative and conditional knowledge (Ratamun & Osman, 2018). However, Hensen et al. (2019) have mentioned that the adverse attitude was due to an instructor effect (i.e. inexperienced teacher-assistants). When this was corrected, they found no significant differences.

In other studies, virtual labs were combined with passive media or hands-on laboratories and were compared with traditional teaching methods alone (only passive media or only hands-on labs). From Table 2.4 we see that these combinations have mostly positive improvements in terms of declarative and conditional knowledge compared to only passive media (2 positive) or only hands-on laboratory (5 positive vs 2 equal). For affective learning outcomes (including attitude, usability and self-efficacy), results are somewhat mixed. On one hand, Astuti et al. (2019) and Kolil et al. (2020) reported positive results in attitude and self-efficacy when virtual labs combined with passive media and hands-on labs are compared with hands-on labs only. On the other hand, Winkelmann et al. (2020) observed no difference in attitude and Enneking et al. (2019) reported that students in the hands-on lab group developed better attitude for laboratory practices than when it is combined with virtual labs. However, when these combined virtual labs are compared with virtual labs only, they seem to be at least as effective in declarative knowledge (2 positive vs 1 equal). Studies with this comparison are uncommon.

Table 2.4: Outcomes of comparative studies including a virtual lab, passive media and/or hands-on lab training. Indications of the following features are given: dk = declarative knowledge; pk = procedural knowledge; ck = conditional knowledge; sb = skill-based; att = attitude; se = self-efficacy; us = usability; o = other; n = number of publications; p = population size.

Compared with	Virtual lab	Passive media	Hands-on lab
Virtual lab	/	**Positive** (n=11, p=1053) references: 11(ck); 14(dk,pk); 33(dk,ck); 34(ck); 35(dk,ck); 40(ck,se); 47(ck,sb,se); 5(dk,sb); 63(o); 70(ck); 61(dk) **Equal** (n=4, p=630) references: 34(dk); 40(dk); 70(dk); 47(dk)	**Positive** (n=6, p=639) references: 1(dk); 53(dk,pk,sb); 60(dk); 64(dk); 65(dk,pk,se,sb); 62(dk) **Equal** (n=9, p=1662) references.: 17(dk,att,se); 24(dk); 30(dk,sb); 32(dk,ck,att); 53(dk,pk,sb); 57(dk,sb); 67(dk,ck,att,us); 68(dk,sb); 64(dk) **Negative** (n=3, p=1260) references: 31(att,us); 32(us); 56(dk,ck)
Virtual lab + hands-on lab	**Positive** (n=1, p=87) reference: 60(dk) **Equal** (n=1, p=141) reference: 1(dk)	/	**Positive** (n=6, p=3060) references: 1(dk); 10(att); 25(dk,pk,ck,sb); 36(ck); 45(se); 49(dk,pk,ck,sb); 60(dk) **Equal** (n=2, p=346) references: 38(dk,ck); 69(dk,att,sb) **Negative** (n=1, p=1141) reference: 25(att)
Virtual lab + passive media	**Positive** (n=1, p=1334) reference: 18(dk)	**Positive** (n = 2, p=360) references: 39(dk,ck,se); 57(dk)	**Positive** (n=2, p=192) references: 10(att); 36(ck)

User studies and performance assessment

Data have been collected from the reviewed publications that evaluated the affective reactions and opinions of the participants on the virtual chemical laboratory. Results were extracted from questionnaires, interviews and observations of the publications where attitude, usability and self-efficacy were examined and are presented in Table 2.5.

The affective learning outcomes of the participant's attitude are separated into two types: attitude towards the subject of chemistry and system usability. The former refers to a person's feelings and beliefs about chemistry and chemical laboratory work, which includes anxiety, satisfaction, intellectual accessibility, usefulness of lab and interest-feeling (Bauer, 2008). Also scientific attitude (Astuti et al., 2019; Chee & Tan, 2012) and open-endedness of lab (Pyatt & Sims, 2012) are added to this category. The system usability refers to the attitude of the participants towards the virtual laboratory application, which includes satisfaction, usefulness of virtual labs, ease of use, time efficiency and realism.

The findings of questionnaires evaluating the usability suggest that participants have overall positive opinions about virtual chemical laboratories. They consider the virtual laboratories to be satisfying, easy to use, useful for learning and take less time than real laboratory work. However, Moozeh et al. (2020) noticed that students were not satisfied by using the virtual lab as post-lab exercise after hands-on laboratory session. In the study of Qvist et al. (2015), difficulties were found in using the user interface and movement control, which affected the students' satisfaction and opinion on time efficiency. Also, some teachers preferred using the real laboratory over the virtual laboratory due to the lack of real laboratory handling and communication between students and assistants (Qvist et al., 2015). In terms of realism, participants found the virtual laboratories to be realistic in studies where 3D, NUI or immersive VR technologies were used. However, other studies reported a lack of realism and authenticity for 2D virtual laboratories (Penn & Ramnarain, 2019; Ramos et al., 2016).

When examining attitude towards chemistry, studies generally revealed positive outcomes in terms of anxiety, interest, usefulness of lab, scientific attitude and open-endedness of labs. Although some of these studies reported no significantly different or worse results compared to hands-on laboratory, participants still displayed positive to neutral attitude towards chemistry and towards usability of virtual laboratories (Enneking et al., 2019; Hensen & Barbera, 2019; Hensen et al., 2020; Winkelmann et al., 2014).

Most studies have reported positive influence on self-efficacy of the participant after using virtual chemical laboratory (Kolil et al., 2020; Makransky, Borre-Gude, & Mayer, 2019). More in particular, virtual labs have increased the participant's confidence in performing laboratory activities and in thinking like a chemist (Ali et al., 2014; Jagodzinski & Wolski, 2014, 2015; Penn & Ramnarain, 2019; Ullah et al., 2016; Woodfield et al., 2005). One study found that confidence was not significantly different from hands-on laboratory. However, they claim that this is because the participants were self-selected and that confidence possibly was increased to the same level as the group of hands-on laboratory (Dalgarno et al., 2009).

Evaluative studies that only assessed the performance of the virtual lab group have measured significant knowledge gain by using knowledge tests (Alqadri, 2018; Annetta et al., 2014) and school grades (Woodfield et al., 2005). Other

Table 2.5: Overview of affective learning outcomes reported in publications that performed user studies and performance assessments of the virtual chemical laboratory applications. The numbers represents the ID number found in Appendix A.2.

Attributes	Sub-category	Positive	Neutral	Negative
Attitude	Attitude towards chemistry	10, 14, 31, 32, 52, 53, 67	17, 25, 69	
Usability	Satisfaction	2, 4, 6, 7, 17, 19, 23, 31, 60, 62, 65, 69, 71, 73, 76	50, 54	
	Usefulness	2, 4-7, 13, 15, 17, 19, 20, 23, 31, 32, 39-41, 50, 52, 53, 55, 57, 60, 62, 65, 68, 69, 71, 73, 76	54	
	Easy to use	2, 4-7, 20, 31, 32, 53, 57, 58, 60, 62, 65, 68, 71, 76		54
	Time efficiency	6, 7, 40, 41, 67, 68, 71, 72, 76	54, 57	
	Realism/ authenticity	2(3D), 5(NUI(3D)), 20(3D), 23(immersive VR), 54(3D), 62(3D)		52(2D), 55(2D)
Self-efficacy	Confidence in laboratory work	5, 39, 40, 45, 47, 52, 61, 65, 71	17	
Performance assessment	Knowledge gain	8, 9, 72		
	Log data assessment	16, 19, 26, 58, 73		

studies were able to distinguish low performing from high performing learners by measuring data during the virtual experience, such as time, number of steps, number of errors and number of hints used (Cuadros et al., 2015; Desai et al., 2017; Gal et al., 2015; Sampaio et al., 2014; Wu et al., 2019).

2.4.2 Technology

Display technologies and natural user interfaces

The type of technology, used in the reported virtual chemical laboratory, was identified by examining indications of the technology in the text and in the images of the publication. We observed two distinct types of technology used in virtual chemical laboratories as explained in Section 2.3.3.

Display technologies are used to visually display chemical experiments or the laboratory environment. The majority of publications have reported the use of virtual laboratories with 3D Desktop technology (n = 37, 49%), while 2D Desktop is second (n = 29, 38%) and immersive VR as third most used (n = 10, 13%). Examples of immersive VR HMD devices used in this study are: Oculus Go, Oculus Rift and Samsung GEAR VR. However, studies using 2D Destop virtual labs seem to have tested more participants (5994 people) than studies with 3D Desktop (5241 people) or immersive VR (511 people). Table 2.6 shows the classification of technology types with the corresponding publications as reference, while Figure 2.5 presents the distribution of the technology types with total population size of the studies per type.

Natural user interfaces (NUIs) can be characterised as input devices to control the virtual environment of the application. NUIs use tracking sensors in order to precisely capture the movement of the user's body. For instance, NUI devices that track the movement or rotation of the hand by using controllers or using cameras such as a Wii Remote or KINECT. With such technology, users can interact with the system in a way that feels more natural up to a point where the user is not aware of the interaction interface between human motion and system control (Jagodzinski & Wolski, 2015). These NUI devices have been used in 5 publications (Ali et al., 2014; Jagodzinski & Wolski, 2014, 2015; Ullah et al., 2016; Wolski & Jagodzinski, 2019). Devices tracking finger gestures, such as Leap Motion controller, have been used in 8 publications (Aldosari & Marocco, 2015, 2016; Al-Khalifa, 2017; Han et al., 2017; Ikram et al., 2015; Y. Kim et al., 2016; Wu et al., 2019). Only one publication reported the use of NUI devices tracking the whole human body such as a KINECT (Desai et al., 2017). As NUIs are used in combination with a visual display technology, publications using 3D Desktop have implemented NUIs (n = 8) more frequently than publications using 2D Desktop (n = 4) and immersive VR (n = 2).

Table 2.6: Number of studies using a type of display technology with or without a type of NUI.

Display	n	References	NUI	N	References
2D Desktop	25	1, 7, 8, 13, 15, 16, 18, 22, 26, 30, 33-35, 38, 45, 49, 50, 52, 53, 55-57, 59, 60, 74	Movement/rotational tracking	3	39, 40, 70
			Hand gestures	1	6
3D Desktop	29	2, 9, 11, 12, 14, 17, 20, 21, 25, 27, 28, 31, 32, 41, 46, 51, 54, 58, 61-64, 67-69, 71, 72, 75, 76	Movement/rotational tracking	2	5, 65
			Hand gestures	5	3, 4, 37, 42, 43
			Body gestures	1	19
Immersive VR	8	10, 23, 24, 36, 44, 47, 48, 66	Hand gestures	2	29, 73

With n = number of studies

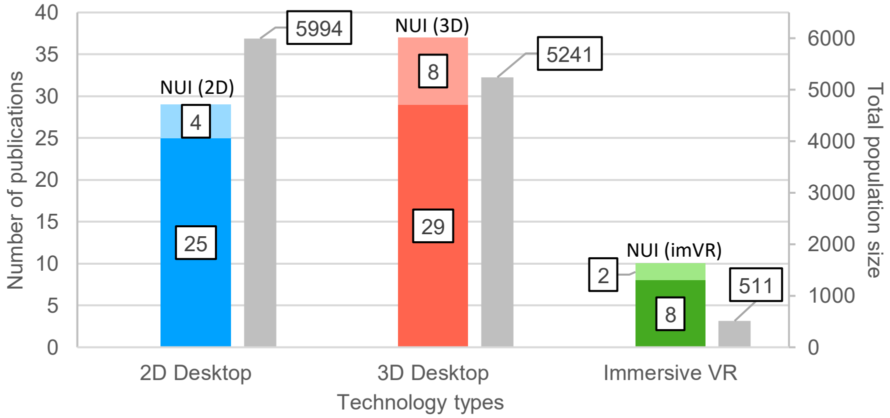


Figure 2.5: Number of publications (left bars) and total population size in all studies (right bars) per technology type. A distinction between display technology without and with NUI (bottom versus top colour) is made

Technology trend

In order to identify the current trend of technology use in virtual chemical laboratories, a distribution of technology types is presented per year over a time span from 2003 to 2020 in Figure 2.6. It shows that research using virtual chemical laboratories with 2D Desktop and 3D Desktop technologies has been prominent throughout the years since the early 2000s with a sharp increase starting from 2012. NUI devices started to be implemented in virtual chemical laboratories from 2014, but have not remained largely present in the last few years. Instead, there has been an increase in the use of immersive technology starting from 2018.

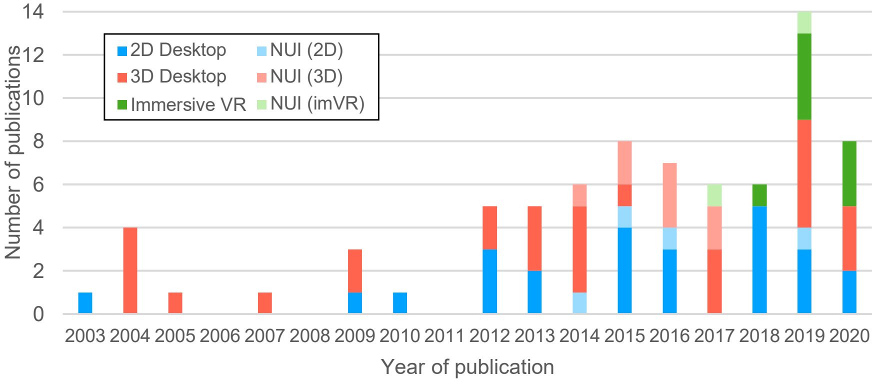


Figure 2.6: Distribution of technology types that are reported in the reviewed publications over a time span of 2003 to 2020.

2.4.3 Instructional design

Learning theories

The learning theories, that are explicitly indicated in the text or as keyword of the reviewed publications, are extracted and presented in Table 2.7 with a short description of each theory. The findings of this review study suggest that inquiry-based learning/discovery learning (n = 7) and learning-by-doing (n = 5) are two of the most used learning theories for virtual chemical laboratories. However, the large majority of the publications (n = 53, 70%) have not specified any learning theory.

Table 2.7: Collection of learning theories used in virtual chemical laboratories of the reviewed publications.

Learning theories	Description	n	Ref.
Inquiry-based & discovery-based learning	A constructivist learning approach where learners are stimulated to conduct an investigation. It requires them to follow the scientific process of formulating questions; hypothesizing the result; investigating and analysing evidences; explaining the findings; and evaluating their arguments (Kebritchi & Hirumi, 2008).	7	13, 14, 22, 26, 53, 57, 65
Kolb's experiential learning	Suggests that knowledge is gained through a four-stage cycle of experiential learning (Kolb, 1984): (1) concrete experience, (2) reflective observation, (3) abstract conceptualization and (4) active experimentation. Following this cycle, the learner encounters a concrete experience that encourages observation and reflection. This reflection of experiences creates abstract concepts which stimulates active experimentation resulting in new experiences. This cycle then repeats itself.	2	20, 61
Learning-by-doing	A form of learner-centred instructional approach that refers to improving skill development of the learner by practical experiences. Learning is, therefore, enforced by experiencing realistic tasks and interaction with the learning environment (Bates, 2015).	5	4, 26, 51, 54, 75
Problem-based learning	Learning approach that introduces challenging problems which the learners needs to solve. It encourages the learner to improve their critical thinking, problem-solving skills and metacognitive knowledge.	2	9, 59
Situated learning	Similar to learning-by-doing and depicts that transfer of learning is improved not only when the learner performs realistic tasks but also when the learner is situated in a realistic environment that is relevant to the learning activities (Wenger, 1991).	2	46, 62
Constructivist-cognitive-contextual	Combination of constructivist, cognitive and contextual learning. Constructivist learning states that learners form their own constructs and adapt their knowledge based on interaction with the surrounding. Cognitivist learning refers to the thinking process when learning is taking place. Contextual learning allows learners to connect content with context by providing real life situations (Bakar et al., 2013).	1	11
Predict-observation-explain	A constructivist learning approach where "students make a prediction and interrogate the nature of situation they faced by combining their existing information with their experiences by using similar situations they faced in real world" (Tatli & Ayas, 2012).	2	63, 64
Not specified		53	

With n = number of studies

Instructional support

Instructional support elements are also extracted from the reviewed publications and are presented in Table 2.8 with the corresponding description. We observed that feedback (n = 11), scaffolding/guidance (n = 8) and modality (n = 6) are the top 3 most frequently used instructional support elements in virtual chemical laboratories. However, similar to learning theories, the large majority of the publications (n = 55, 72%) have not specified any instructional support.

Value added research

In order to evaluate the effectiveness of some of the instructional support elements, a small part of the comparative studies have performed value-added research. Table 2.9 shows an overview of these comparisons and the achieved results. From these studies, the instructional support principles of modality and spatial contiguity have been confirmed when applied to virtual chemical laboratories. Providing instructions in audio voice is observed to be more effective than textual or video instructions (Jagodzinski & Wolski, 2015) and displaying learning information near the object seems to be more effective than when it is unrelated to its position (Zayas-Pérez & Cox, 2009).

Two studies have examined the scaffolding/guidance principle. Ullah et al. (2016) found that giving procedural guidance (i.e. step-by-step instructions) resulted in better procedural knowledge compared to no guidance by measuring time of completion and errors made. However, it seemed that the participants score equally in declarative knowledge and skill-based outcome by test and hands-on lab practical evaluation. Borek et al. (2009) studied the effect of offering minimal guidance (i.e. inquiry), guidance when needed (i.e. tutored) or explicit instruction (i.e. direct instruction). They found that tutored approach resulted in better conditional knowledge than inquiry and direct instruction approaches. However, the results were not significantly different for declarative knowledge. They suggested that learners need sufficient guidance while using the virtual chemical laboratory, but not too much as they are demotivated by the lack of autonomous decision-making.

Also, a study on pedagogical agents was performed by Makransky et al. (Makransky, Terkildsen, & Mayer, 2019). They examined the effect of the virtual agent's appearance on the learning performance of boys versus girls between the ages of 13 and 16. A version of the commercial virtual laboratory Labster was altered to display a drone guide while another version used a human female guide. Results showed that boys performed better with the drone guide and girls performed better with the human female guide. The study concluded that gender matching of the pedagogical agent could motivate the learner to do more effort to learn.

Table 2.8: Collection of instructional support used in virtual chemical laboratories of the reviewed publications.

Instructional support	Description	n	Ref.
Feedback	Information that is provided by an agent (e.g. teacher, book, self, game, etc.) as a consequence of the performance of the learner (Hattie & Timperley, 2007). It has the purpose to direct learners to evaluate their progress towards a goal, identify knowledge gaps and reduce discrepancies between current understanding and the intended goal (Johnson et al., 2017).	11	13, 19, 22, 24, 26, 44, 47, 48, 50, 61, 75
Scaffolding & guidance	An instructional technique that provides guidance and instructional support to the learner in an adaptive way (Pea, 2004).	8	13, 18, 19, 22, 26, 44, 61, 65
Pedagogical agent	The pedagogical agent, as Martha and Santoso (2019) have stated, "... is an agent (single or multi) in the form of a virtual character equipped with artificial intelligence that can support the students' learning process and various instructional strategies in an interactive learning environment".	4	9, 24, 48, 51
Modality	A multimedia principle which states that it is more effective when information is given by a mixed modality presentation (partly visual and partly auditory) than only one modality type (either visual or auditory) (Low & Sweller, 2014).	6	5, 15, 47, 48, 65, 75
Personalisation	A multimedia learning principle which explains that people learn more deeply when written or verbal presentations are in conversational style rather than formal style (Mayer, 2014b).	2	15, 48
Narrative element	Narrative depicts the use of a story as a teaching tool that allows the learner to construct a cognitive framework to structure the information and experiences (Wouters & van Oostendorp, 2013).	2	9, 14
Reflection	Reflection is the meta-cognitive process of the learners where they reflect on their own learning process and the decisions they make during this process (Flavell, 1979).	1	47
Spatial contiguity	A multimedia principle which states that deep learning of the learner is better achieved when visual information (e.g. texts, pictures, animations) is presented near the relevant learning content rather than far away (Mayer & Fiorella, 2014).	1	75
Not specified		55	

With n = number of studies

Table 2.9: Comparative studies with value-added research and the achieved results.

Instructional support	Comparison	Results	Ref.
Scaffolding & guidance	procedural vs no procedural guidance	procedural $>$ no procedural guidance (pk) procedural $=$ no procedural guidance (dk, sb)	65
	inquiry vs tutored vs direct instructions	tutored $>$ inquiry $=$ direct instruction (ck) tutored $=$ inquiry $=$ direct instructions (dk)	13
Pedagogical agent	robot drone vs human female	human female $>$ robot drone (for girls) (dk, ck) robot drone $>$ human female (for boys) (dk, ck)	48
Modality	voice vs text vs video instructions	voice $>$ text $=$ video instructions (dk, ck)	39
Spatial contiguity	co-located vs not co-located information	co-located $>$ not co-located (dk)	75

With dk = declarative knowledge; pk = procedural knowledge; ck = conditional knowledge; sb = skill-based.

2.5 Discussion

2.5.1 Research on virtual chemical laboratories

In this review study it was found that most publications performed media comparative studies in order to compare the performance between virtual labs and traditional teaching methods. Mainly quantitative evaluation methods were used for comparison, such as knowledge tests to examine the cognitive learning outcomes and lab practical assessments to assess practical laboratory skills. Similar to other reviews about virtual laboratories, declarative knowledge is the most studied learning outcome in this study (Brinson, 2015; Ma & Nickerson, 2006). Additionally, qualitative evaluation methods, such as questionnaires, interviews and observations, were also used in comparative studies to perform user studies.

Results of media comparison studies reveal that the effectiveness of virtual chemical laboratories varies widely depending on which traditional teaching method they are compared with. When compared with passive media (e.g. classroom lectures, text or video), virtual labs are more effective for improving conditional knowledge but are, for some studies, not significantly different in terms of declarative knowledge. This means that for learning basic knowledge of chemistry facts and concepts, virtual labs are sometimes equal to passive media. This could be seen in the study of Makransky et al. (2019) where this was explained by the cognitive overload (i.e. overwhelming cognitive capacity) of the learner while using VR systems. However, virtual labs do show better results when learners need to reason and apply chemical concepts to solve problems (Herga et al., 2015; Jagodzinski & Wolski, 2015; Makransky, Borre-Gude, & Mayer, 2019). Virtual labs are able to provide dynamic visualisations in the sub-microscopic domain, while also offering an interactive platform for the learners (Herga et al., 2015). This combination of visual support and high level of interactivity engages the learner to develop a deeper understanding of the learning content (Davenport et al., 2018; Trindade et al., 2002). Combining virtual labs with passive media seems to result in a greater improvement as it reinforces the previously learned concepts (Davenport et al., 2018).

Different results are found when virtual chemical laboratories are compared with traditional hands-on laboratories. These comparative studies suggest that virtual chemical laboratories are equally effective or sometimes better than hands-on laboratories regarding declarative knowledge, procedural knowledge and skill-based outcomes. These findings align with outcomes of other literature reviews in which also equal or improved results were observed between non-traditional (e.g. virtual, remote and at-home kit) and traditional laboratories (Brinson, 2015; Sypsas & Kalles, 2018). While it is frequently argued that virtual labs cannot replace real hands-on laboratories (Ikram et al., 2015; Penn & Ramnarain, 2019; Zhong & Liu, 2014), very little evidence has been found that virtual labs perform worse than hands-on labs (Faulconer & Gruss, 2018). This means that learners do learn procedural knowledge and laboratory skills in virtual environments where physical interaction is limited (Pyatt & Sims, 2012). Especially when procedural guidance was provided during the virtual experiment, learners were able to perform better than their peers who were trained in the real laboratory (Ullah et al., 2016). However, it is possible that the lab practical experiments were so simple that simple interactions in virtual labs are sufficient to learn the techniques (Winkelmann et al., 2017). More research is required to investigate practical laboratory skills in virtual labs as there is a lack in studies that assess skill-based learning outcomes. So despite the media comparison of virtual and real labs steers towards equal effectiveness, virtual labs still have the advantage that no physical lab environment is needed, thus reducing cost, time, staff personnel and allowing easy accessibility (Brinson, 2015). Furthermore, a more effective use of virtual laboratories is to utilise them as a supplementary tool combined with hands-on laboratory resulting in improved cognitive and skill-based outcomes. When virtual labs are provided as pre-laboratory exercise, self-efficacy of students was significantly improved compared to hands-on lab only (Kolil et al., 2020). However, one must be careful not to overwhelm students with extra work load (Woodfield et al., 2005) or demotivate them with post-laboratory exercises (Moozeh et al., 2020).

Compelling results are found when attitude and usability towards chemistry laboratory are compared. Some studies show no difference in attitude (Dalgarno et al., 2009; Hensen et al., 2020; Winkelmann et al., 2014, 2020), while other studies observed worse results when compared to hands-on laboratories (Enneking et al., 2019; Hensen et al., 2020). Some students seem to believe traditional labs are more useful and easier to use than the virtual labs (Enneking et al., 2019; Hensen et al., 2020). The reasons for these findings are still unclear, but authors suggest it could be due to self-selection bias (Dalgarno et al., 2009) or instructor effect (Hensen & Barbera, 2019). It can also be noticed that the studies of Enneking et al. (2019) and Hensen et al. (2020) have used the same virtual lab called LearnSmart Laboratories. So the discrepancies in attitude can be affected by the design of the virtual application. Nevertheless, comparison of affective measures between different media should be more rigorous to minimise any kind of bias.

Evaluative studies are the second most common research purpose in this review. These studies only considered the group using virtual chemical laboratories in order to evaluate the affective learning outcomes of the participants with questionnaires as the most used evaluation method followed by interviews and observations. The results of these user studies reveal, in general, positive attitude towards chemistry, good usability of the virtual lab and improved perceived self-efficacy; despite that some studies reported significant better results for hands-on laboratory. This means that most users in the included publications consider the virtual laboratories to be satisfying, easy to use, helpful for learning and take less time than real laboratory work. These positive reactions and opinions indicate that the students and teachers accept to use these systems as educational tool for laboratory practices. However, as stated previously, it depends on the design and implementation of each individual virtual laboratory. Usability issues should be resolved (Qvist et al., 2015) and teachers should be well trained in using these applications (Hensen & Barbera, 2019) in order to provide better experiences.

Other evaluative studies demonstrated the possibility to evaluate procedural knowledge and skill-based outcomes by utilising real-time assessment during the virtual experience (Cuadros et al., 2015; Desai et al., 2017; Gal et al., 2015; Sampaio et al., 2014; Wu et al., 2019). This opens doors of opportunities for unintrusive transfer tests that could reduce test anxiety as the student is unaware of the assessment (Shute, 2011) and could avoid replication of real-life lab practical tests. However, there is a lack of studies using this evaluation methodology. Furthermore, this stealth assessment raises ethical and social issues by evaluating students without them being aware of the evaluation (Georgiadis et al., 2018). Therefore, when using this assessment methods, students must be well informed and consent to such practice.

2.5.2 Technology use in virtual chemical laboratories

The technologies used for virtual chemical laboratories in this review study are distinguished in visual display output technology and kinaesthetic NUI input devices, where display technology are further divided in 2D and 3D graphics on monitor displays, and immersive VR headsets. This technology distinction is similar to the work of Ali and Ullah (2020), however in this case, immersive VR and NUI are added because these innovative technologies are distinct from 2D and 3D Desktop in a way that they have the capability to simulate the chemical laboratory more realistically.

Virtual chemical laboratories with 2D Desktop technology have been used primarily to provide simple dynamic visualisation and simulation of chemical experiments. They can display easy comprehendible animations that integrates the three levels of chemical representation (Johnstone, 1991): macroscopic (e.g. colour, solid, liquid), sub-microscopic (e.g. atoms and molecules) and symbolic level (e.g. chemical notation). With these animations, they support the learner's understanding of chemical reactions at sub-microscopic level, offering an advantage over traditional media (Herga et al., 2015). Moreover, free experimentation is possible without requiring a real laboratory environment (Yaron et al., 2010). However, one of the drawbacks is that 2D representations are unable to provide realistic laboratory environments and actual lab skills (Ali & Ullah, 2020; Penn & Ramnarain, 2019). Despite this lack of realism, they have been used consistently over the years with large population sizes. A reason for this could be that the simple geometries allow easy implementation via internet as they are less demanding in terms of computer performance and internet bandwidth than more advanced 3D VR systems (Ali & Ullah, 2020).

A majority of publications have used 3D Desktop technology. These virtual chemical laboratories were developed with more realistic and more accurate representations of laboratory environments (e.g. fume hood, lab benches, cupboards with chemicals) and laboratory equipment (e.g. flasks, burettes, pipettes) than 2D Desktop laboratories. Additionally, users were also able to explore the simulated laboratory and freely manipulate 3D objects (Dalgarno et al., 2009; Qvist et al., 2015; Winkelmann et al., 2017; Zayas-Pérez & Cox, 2009). Authors agree that this level of realism and interactivity of virtual chemical laboratories can help students to familiarise with the laboratory prior to real laboratory practices (Dalgarno et al., 2009; Georgiou et al., 2007; Tarng et al., 2017). Another use of realistic simulations is the possibility to simulate hazardous events that would otherwise be too dangerous to experience in real life. As such, unsafe laboratory handling can be recognised and good laboratory practices can be taught in the virtual environment without putting students at real risks (Bell & Fogler, 2004; N. S. Dholakiya et al., 2019; Makransky, Borre-Gude, & Mayer, 2019). However, virtual chemical laboratories in a 3D environment require more computing power due to the cost of rendering 3D objects with multiple polygons (i.e. geometries that a 3D object is made of) and interactions of the users simultaneously in real-time (N. D. Dholakiya et al., 2019; Tarng et al., 2017). Nowadays, the huge improvement of recent computer technologies have made it easier to realise virtual environments with this high level of realism and interactivity, unlike in the early years of the computer age (Dalgarno & Lee, 2009). Still, 3D virtual chemical laboratories, that are displayed on a computer monitor using keyboard and mouse, are unable to bring the same feeling and practical handling as laboratories in reality (Dalgarno et al., 2009; Winkelmann et al., 2017).

Recently, immersive VR technology has been emerging as a promising educational tool for virtual chemical laboratories. With HMD VR devices, this technology offers a high level of immersion providing the feeling of really ‘being there’ in a virtual laboratory environment, whereas 3D Desktop is considered only as a low immersion technology because of the external screen (Buttussi & Chittaro, 2018). The technological advancement of 3D stereoscopic depth, head position/rotation tracking and visual isolation from the real world makes the user believe that he or she is in an actual laboratory, thus taking a closer step to virtually replicating a chemical laboratory with high realism. Also, it is believed that the increased motivation and engagement positively influences the cognitive learning outcomes (Pekrun, 2000). However, when comparing an immersive VR virtual lab with passive media and hands-on laboratory, studies reported equal effectiveness in declarative knowledge (Dunnagan et al., 2020; Makransky, Borre-Gude, & Mayer, 2019). Other drawbacks are: more expensive than 2D and 3D Desktop; possibility to induce simulator sickness; and social isolation (Fransson et al., 2020; Pohl & de Tejada Quemada, 2016). While immersive VR might not be the most efficient tool to teach declarative knowledge, perhaps it has a better use as behavioural and emotional training tool in certain laboratory situations (Makransky, Borre-Gude, & Mayer, 2019).

In addition, visual display output technology can be combined with NUI input devices in order to enhance the physical authenticity of the virtual chemical laboratory. With NUI technology, it is possible to have realistic interactions (e.g. grabbing, pinching, pouring) with virtual objects in an ergonomic way due to advanced tracking technology (Al-Khalifa, 2017; Jagodzinski & Wolski, 2015). Using the KINECT technology, it is possible to further increase the sense of presence and immersion by positioning the user's body within the virtual laboratory environment (Desai et al., 2017). Some studies in this review already have successfully implemented this technology to perform chemical experiments (Aldosari & Marocco, 2016; Han et al., 2017; Wu et al., 2019). However, there are some limitations yet to be overcome such as, inability to precisely capture fine hand gestures, and to touch or smell the virtual objects (Jagodzinski & Wolski, 2014; Wu et al., 2019). Nevertheless, combination of NUI technology with immersive VR devices promises great opportunities to exactly replicate real-life chemical laboratories and the practical skills in a virtual environment (Wu et al., 2019).

2.5.3 Instructional design of virtual chemical laboratories

In this systematic literature review, learning theories and instructional support elements were investigated that have been implemented in virtual chemical laboratories. There is a common argument in literature that learning theories are often neglected in studies of educational technology (Hew et al., 2019). Especially with the use of VR technologies, integration of instructional design features is a necessity (Makransky, Borre-Gude, & Mayer, 2019). Unfortunately, the findings in our literature study could not disprove this argument, as a majority of the reviewed publications did not specify a learning theory. Learning theories are important because they can describe, explain and predict how people learn when using certain technologies (Hew et al., 2019). In this way, instructional design of virtual chemical laboratories can be adapted to these theories to maximise learning mechanisms. The studies in this review that did specify learning theories, have most frequently mentioned inquiry-based learning, discovery learning, learning-by-doing and experiential learning. This is not surprising because inquiry and discovery learning are inherent characteristics of laboratory instructions (Domin, 1999). The interactivity and autonomous learning are aspects of virtual chemical laboratories that make these learning environments constructivist and learner-centred. This allows the learner to create a more meaningful understanding of chemical concepts (Tatli & Ayas, 2012).

Another aspect of instructional design is the instructional support that the learner receives during the virtual learning experience. As seen in the study of Makransky et al. (2019), effective learning in virtual environments can be hindered by cognitive overload of the learner. Therefore, providing instructional support could manage this cognitive load more efficiently and could assist the learner when needed (Wouters & van Oostendorp, 2013). Instructional support elements, such as feedback, scaffolding/guidance and modality, have been used the most in the studies of this review. Although most of these studies have only briefly mentioned these features, more can be learned from studies that have performed value-added research on instructional support principles. These studies have investigated the effectiveness of certain types of instructional support by comparing different versions of the virtual application. According to these studies, it is suggested that virtual chemical laboratories are most effective when instruction is given near the location of the learning content (i.e. spatial-contiguity principle) using audio source (i.e. modality principle) and when guidance (e.g. procedural instruction, hints, feedback) is given only when needed (Borek et al., 2009; Jagodzinski & Wolski, 2015; Ullah et al., 2016; Zayas-Pérez & Cox, 2009). However, not enough studies have conducted value-added research in the context of virtual chemical laboratories. Also, similar to learning theories, a majority of the studies have not specified any instructional support element. Eventually, we have come to a point where it seems that we should focus more on how virtual chemical laboratories are designed rather than merely comparing different instructional media (Hu et al., 2022). In this way, we can find a more meaningful progress in research leading to more effective virtual laboratory systems.

2.6 Conclusions

This literature review shows an analysis of published research that has been done on virtual laboratories for chemistry education. The current review adds on previous reviews in this field because we focused not only on the effectiveness of virtual labs in chemistry education but also included an in-depth analysis on both novel technology and instructional design.

The results of this review conclude that virtual chemical laboratories are viable as an effective complementary tool or as an alternative to hands-on laboratories, despite several publications have argued that they cannot be used as replacement (Ikram et al., 2015; Penn & Ramnarain, 2019; Sypsas & Kalles, 2018; Zhong & Liu, 2014). Virtual labs can provide better results in learning outcomes of all domains (i.e. cognitive, affective and skill-based) than traditional passive media and they are considered to be equally as effective and sometimes better than real hands-on laboratories. A more effective use is to combine virtual labs with passive media or with hands-on labs. However, important considerations need to be taken in terms of choice of technology and instructional design.

Technologies used in virtual chemical laboratories range from simple 2D graphics to more sophisticated 3D representations of the real laboratory. Even though 3D Desktop has been used more than 2D Desktop and immersive VR, each of these technologies have their own benefits and have different purposes. One might opt for a low-cost easy to implement 2D virtual labs to teach chemical reactions, or a more costly complex 3D virtual lab to replicate experiments with simple interactions. If high realism is required, the more expensive immersive VR technology and NUI input devices can be used.

This review also found that most studies have not considered learning theories or instructional support in the instructional design. However, these elements are essential to efficiently manage the learner's cognitive load and provide sufficient assistance when learners are struggling.

This literature review can be helpful for researchers, teachers and instructional developers to implement effective technologies and instructional design elements that are based on research on virtual chemical laboratories. Even though virtual laboratories cannot provide the real experience and skills as real laboratories with current technology, they are still effective tools for distance learning. Especially for situations when distance learning is the only option, such as in pandemic outbreaks, schools that cannot afford the cost of real laboratories or individuals who are unable to attend certain laboratory sessions.

There are some limitations of this literature review. Firstly, there is a possibility that we have overlooked an unknown number of publications that could be included in this review due to: only one database was used (i.e. Web of Science); and some publications were excluded as they did not clearly specify a chemical laboratory practice or technology. Nevertheless, this limitation should not have affected our conclusions severely. Secondly, the data of comparative and evaluative studies were not compared quantitatively in detail. In order to know the effect size of how much the effectiveness is of virtual chemical laboratories, a systematic meta-analysis is needed. Thirdly, other immersive technologies, such as Cave Automatic Virtual Environment (CAVE) and Augmented Reality (AR) were not included because they are not fully virtual and still require a physical space in the real world. Fourth, one should be taken into account that a systematic literature review only reports on the choices made in publications and on the published results which are sometimes presented without a full background reasoning. For example, the collected instructional designs and theories are based on publications mentioning the method, but no information is given on why these choices were taken. Finally, positive outcomes of this systematic literature review might be elevated by publication biases. For example, the lack of publications with negative results due to investigators simply not reporting findings with negative results.

In this chapter, there is no in-depth analysis of virtual laboratories for safety training, because there is a gap in literature research about this topic. Lab safety training can benefit from virtual laboratories since these can be ideal training environments for laboratory workers to practice lab safety procedures without being exposed to real chemical hazards. Following the learning theories of situated learning and learning-by-doing , learners can potentially learn lab safety skills better when they experience and safely handle dangerous situations simulated in virtual labs (Miliszewska & Sztendur, 2011). Immersive VR technology is the ideal tool to let learners be fully immersed in such situations and enact safety measures with realistic interactions. In addition, the high level of interactivity could help the learner be more engaged with safety training, which are otherwise unmotivating to follow with conventional training methods. For these reasons, there is a need to investigate virtual laboratories using immersive VR technology as a tool to motivate and engage learners to follow lab safety training.

CHAPTER 3

Design and Development of a VR Serious Game for Chemical Laboratory Safety

In order investigate motivation and engagement of VR lab safety training, such VR tool should be optimally designed to increase motivation. Digital games are known for such positive effects. Therefore, this chapter fulfils this need by describing the design and development process of a VR serious game for chemical lab safety training, called VR LaboSafe Game. First, the design of the VR game is described with consideration of well-known design principles. Then, the development process of the VR LaboSafe Game is reported with the use of easily accessible game development software tools. Although the VR LaboSafe Game is presented as an example, the proposed design guidelines and development tools are generic and can be applied to other VR safety training development projects.

This chapter is adapted from : Chan P., Van Gerven T., Dubois J.-L., Bernaerts K. (2021). Design and Development of a VR Serious Game for Chemical Laboratory Safety. 10th International Conference on Games and Learning Alliance (GALA), Online. Springer International Publishing, Cham. 1: 23-33.
Author's contribution: Chan P. designed and developed the VR LaboSafe Game, and drafted the manuscript. Bernaerts K., Dubois J.-L. and Van Gerven T. managed and supervised the project.

3.1 Introduction

Since the beginning of the computer era, video games have become more and more integrated in modern society. This can be noticed by its increasing popularity, where nearly 40% of the world population are playing video games, according to a report in 2021 (DFC Intelligence Research, 2021). The revenue of global video game industry even surpasses the global film and music industry combined (Ward, 2021). While the interest of games for entertainment increases, also the interest of games for other purposes increases, including games for education and training – also known as serious games. The benefits of games are not only known to have positive effects on motivation and engagement, but also encourage positive learning outcomes (Garris et al., 2002; Gloria et al., 2014).

Along with this popularity increase and together with the recent technological advancements of computers, multiple software tools have emerged to make game development easier and more accessible to a wider public. Game engines with a free licensing option, like Unity, Unreal Engine and Godot, have become more advanced in a way that even people with limited programming skills and knowledge are able to develop their own game (Gloria et al., 2014).

Although developing your own game has become more accessible, the process of design and development of a serious game for education and training is still a difficult process, mainly because it requires a strong collaboration of experts in many different fields. A typical multidisciplinary team consists of educators, instructional designers, game developers, visual artists, scriptwriters and many more (Dimitriadou et al., 2020; Garcia Fracaro et al., 2021). Nowadays, it is possible to learn most of the development on your own via self-directed learning with online resources, but the large amount of different design principles in scientific literature can make it confusing for a beginner to know where to start.

3.2 Designing a VR serious game for chemical lab safety

The Design phase, in an instructional design process, is dedicated to define learning tasks and testing strategies (Branch, 2009). From the perspective of a game life cycle, this phase corresponds to the pre-production phase where a game design is created that elaborates the gameplay and game mechanics (Dimitriadou et al., 2020; Ramadan & Widyani, 2013). However, it is not an easy task to successfully align learning tasks and game mechanics in serious games.

Furthermore, there are some other complications for designing effective VR learning experiences. For example, poorly designed VR serious games can cause users to feel nauseous, present an overwhelming amount of information or can simply not be motivating to play (Kourtesis et al., 2019; Makransky, Terkildsen, & Mayer, 2019; Hu et al., 2022). So, it is a more complex interplay between cognitive capabilities and psychological factors of the learner.

In general, designing such complex training systems is not easy and requires many factors to be considered in order to maximise its effectiveness. Therefore, this section provides several well-researched design principles that could overcome these challenges. Examples are given how these principles are implemented in the game design of VR LaboSafe Game. These design choices that are implemented in the VR serious game are not only based on design principles from literature, but are also inspired from the gaming experience of the designer (i.e. the author of this dissertation).

3.2.1 Combining instructional design and game design

In order to design a serious game meant for educational purposes, the game should be designed with a solid instructional design, as well as an engaging game design. However, the challenge is to find the optimal balance between the two designs. Serious games that are heavily focused on transfer of learning might not be fun to play, while other games designed to be entertaining might not be educational (Arnab et al., 2015). Ideally, instructional design and game design should be complementary to each other. Therefore, we need to first determine what the key elements are for what defines an instructional design and for what defines a game design.

Instructional design elements

The instructional design is the main structure of the learning experience in a serious game. It describes what learning objectives the learners will have to achieve, how the learning content can be presented and how the performance of the learner will be assessed.

The learning objectives of an instructional intervention determine what knowledge, skills and attitude are expected of the learner to have gained after completing the learning experience. These learning objectives provide the initial framework which the game is built upon (Schrier, 2014). They have a large influence on how the game looks and what activities the player would be performing to achieve the intended learning outcomes. Because of its importance for the design of the whole game, learning objectives should be precisely defined at first before continuing to consider other design elements. From an instructional point of view, it is widely known that effective learning objectives need to be clear, reachable and measurable (Weitze, 2014).

After clearly defining the learning objectives, an *instructional approach* is chosen that describes how the learning content is presented and how the learning activities are aligned with the learning objectives. Serious games commonly apply student-centred approaches so that the learners have more control over their own learning process which makes the learning activities more interactive (Kebritchi & Hirumi, 2008). Ideally, these learning activities should coincide with the game activities in the serious game. Arnab et al. (2015) proposed a model how to align learning mechanics with game mechanics and Rapeepisarn et al. (2008) presented a relationship between game genres, learning techniques and learning styles.

Another important element in the instructional design is *the assessment method.* It refers to measuring and using data to demonstrate that the learners have actually accomplished the learning objectives and how effective they have performed the learning tasks. This assessment can be taken in a summative way (e.g. with a test or survey at the end of a performance) and/or in a formative way (e.g. with data logs of the player's performance during gameplay) (Bellotti et al., 2013). The latter, also known as in-game assessment, creates the opportunity to assess the learners' complex cognitive skills, such as problem solving and reasoning skills, by tracking their actions within the game (Udeozor et al., 2021). In order to avoid breaking the flow of the game, the assessment should be integrated in a way that is less intrusive and less obvious for the player (Shute, 2011).

Game design key elements

The game design determines the main structure of the game experience of a serious game and consists of several game characteristics. Numerous scholars have different definitions of what a game is and what the main characteristics are that defines a game (Garris et al., 2002; Prensky, 2001a; Salen & Zimmerman, 2004). Usually, it comes down to these elements: goal, rules and challenge (Wouters et al., 2013). Defining these core elements gives a base structure of the serious game design and specifies what activities the player will be doing during the game.

The goal of a game is the central feature of its formal structure of the game as a whole that is defined by the game rules, according to Salen and Zimmerman (2004). They have suggested that games should have clear and quantifiable goals, otherwise players cannot judge if their actions will reach the goal, making it impossible to have meaningful gameplay. When designing serious games with the main purpose to achieve a learning outcome, the game goals should be designed so that they support the learning objectives to be taught.

Rules have a purpose to impose limits and to force the player to select a specific path to reach the goals of the game while ensuring that every player adheres to the same rules (Prensky, 2001a). These rules dictate when the player is in a winning state or in a losing state. In other words, rules represent criteria to evaluate the player's performance and progress in the game in the form of scoring or progress bars (Huang & Johnson, 2009). Therefore, in the context of serious games for learning, rules are not only a way to determine if the goals of the game are reached, but also if the learning objectives are reached.

The challenge in a game is another important aspect in a game design. It presents a level of difficulty to achieve the game goals, which maintains the engagement of the player in the game and is a driving force to reach these goals (Malone, 1981). Optimal challenges should be not too difficult in a way that causes anxiety, nor too easy in a way that increases boredom. In doing so, 'a state of flow' can be achieved within the player, which depicts "state in which people are so involved in an activity that nothing else seems to matter" (Csikszentmihalyi et al., 1990). Thus, when serious games for learning purposes control and maintain the flow of the learners, enhanced engagement and attention on the learning material can be accomplished. To maximise learning, challenges should be designed that encompass the learning goals.

3.2.2 Determining the learning objectives

Before designing the VR serious game, the learning objectives of the chemical laboratory safety training are determined. The target audience consists of people who frequently work in a chemical laboratory, including students, researchers and lab technicians. In order for them to work safely in a laboratory, they should be able to demonstrate the four main principles of lab safety skills, known as RAMP (Hill, 2010): Recognize hazards, Assess risks, Minimize risks, and Prepare for emergencies. Therefore, the main objective of a lab safety training is to improve the trainees' safety awareness and safety behaviour. Safety awareness is the constant consciousness of the state of safety in the surrounding environment, while safety behaviour, in our case, refers to the practical application of safety measures in the laboratory.

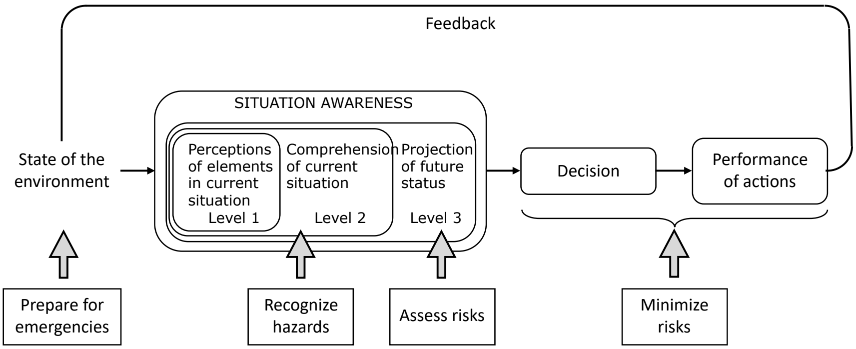


Figure 3.1: Diagram of the human decision-making model of Endsley (Endsley, 1995) adapted with the RAMP principles of Hill et al. (Hill, 2010).

In context of human cognitive processing in a given situation, these principles could be aligned with *the human decision-making model* of Endsley (1995), where situation awareness refers to safety awareness and the decision-making process refers to safety behaviour. A schematic overview is shown in Figure 3.1 how people process information from a situation to make decisions. This means that people working in high-risk environments should perceive and comprehend the hazards in their surroundings, predict possible safety risks that can happen and make decisions to minimise these safety risks. In case the state of the environment gets worse, they should make further decisions to prepare for emergencies.

While adapting these principles and using the *revised Bloom's taxonomy learning verbs* (Anderson et al., 2018), learning objectives of the VR LaboSafe Game are defined in order to train lab workers to make safe decisions in a virtual chemical laboratory. This means that after playing the VR LaboSafe Game, the trainee is able to:

- **Understand** and **recognise** chemical hazards that are present In the chemical laboratory (i.e. hazard identification)
- **Assess risks** in the chemical laboratory and **predict** how likely a hazard can cause an accident in order to make correct decisions for effective safety measures (i.e. risk assessment)
- **Determine** appropriate safety measures and **apply** them to eliminate or minimise risks in the chemical laboratory (i.e. risk management)

- **Demonstrate** safe procedural skills of laboratory procedures and **conduct** a chemical experiment safely (i.e. safe laboratory procedure)

3.2.3 VR LaboSafe Game design

VR LaboSafe Game is a serious game that utilises immersive virtual reality for chemical laboratory safety training. In this section we describe the gameplay and how the game design is aligned with its instructional design.

General game overview

The genre of the game is a single-player simulation game with a problem solving characteristic. A realistically accurate laboratory environment is simulated with task-based activities that are related to laboratory activities in real-life (See Figure 3.2). Using *the situated learning* approach, the trainee learns about laboratory safety by physically doing activities related to the skills of hazard identification, risk assessment, risk management and safe laboratory procedures. The main goal of the game is to complete these objectives with a game rule to not cause accidents that will deplete the player's health points. Throughout the game, there is a friendly robot in the shape of a chemical batch reactor, called Archy, that guides the player through the different game levels and provides feedback on the player's performance.

Level design

The VR LaboSafe Game has three different game modes: tutorial mode, training mode and evaluation mode.

The tutorial mode has the purpose to get the player familiarised with the controls and interactions of the game (See Figure 3.2). Especially for beginners who are new to VR, tutorial sessions are recommended prior to using the game to its full extent (Checa & Bustillo, 2019). In this way, the player could be able to handle the controls more easily during the learning experience, which leads to performing the tasks more effectively. We developed this tutorial in separate sections. The first tutorial section includes instructions for interacting with Graphical User Interfaces (GUI) in the 3D environment, teleporting in the environment, grabbing objects and pausing the game. In the second tutorial, players learn how to use a virtual tablet that they can activate at any time. This tablet with a virtual touch interaction can be used to display the game objectives and game progress, scan chemical containers to obtain the Safety Data Sheet (SDS) documents of the chemical, and take pictures of the virtual environment. The third tutorial introduces the player to the health points system and how to protect their health points by equipping Personal Protective Equipment (PPE) clothing. Because these VR interactions can be quite complex for first-time users, each instruction has small minigame tasks (e.g. collecting all balls in time) in order for the player to be acquainted to interacting and handling in VR by *part-task practice* (Kirschner & van Merriënboer, 2018).

The training mode has the purpose to train the player's skills of laboratory safety in an interactive and engaging way. There are three distinct game levels, also called missions, with different tasks: 1) Risk Spotting; 2) Risk Minimisation; and 3) Safe Experiment. The first mission level is designed with a 'search and find' game mechanic, where the player has to spot a certain number of safety risks that appear in the virtual laboratory (See Figure 3.2). For example, some safety risks are: a lab technician working in a fume hood with the window fully open; a bottle with flammable chemical is stored in a fridge that is not safe against ignition; or a lab technician that is not wearing the correct PPE. The player takes a picture of these risks and answers questions on the laboratory tablet. These questions are related to the chemical hazards and consequences of the spotted risks in order to train and assess the skills of hazard identification and risk assessment. This game level is completed when all risks are found. The second mission level is an extension of the first game level in a way that the player not only needs to spot risks in the lab but also needs to correctly eliminate or minimise these risks (See Figure 3.2). For example, when the player spots a full beaker of nitric acid waste, then the player should dispose this chemical waste in the correct waste container. This level develops the player's risk management skills. In the first and second level, the challenge can be tailored by adjusting the number of risks and their complexity. The third mission level is distinct from the first two levels in which the player needs to complete a chemical experimental procedure with the necessary safety measures (See Figure 3.2). For example, the player needs to perform a distillation experiment of ethyl acrylate as safe as possible. Unsafe and dangerous decisions of the player can result in an accident with a reduction of health points. This level allows the player to train the procedural skill of managing a chemical experiment safely. The challenge in this level depends on the complexity of the experiment setup.

The evaluation mode is where the safety awareness skills and safety behaviour of the player are taken to the test. The different objectives, that are also found in the training levels, are combined into one game level where the choices of the player can influence the scenario of the level. For example, the player needs to spot safety risks in the beginning of the level before performing an experimental procedure. When crucial safety risks are not recognised and eliminated, these risks can cause an accident during the chemical experiment procedure. The in-game assessment of the player's decisions and actions can evaluate the player's competence in laboratory safety.

In-game assessment During every mission level, in-game measurements caused by the player's actions are collected and recorded. Game elements that are saved in a data log file include: number of correct or incorrect actions, total time of finishing a mission level, number of hints requested, health points remaining and protection points lost. At the end of the level, these measurements are analysed to assess the players' performance on a three-star rating system: Novice, Regular and Expert. Because each mission level is designed based on learning objectives, the players' game performance could give an indication of their safety competence in a real laboratory.

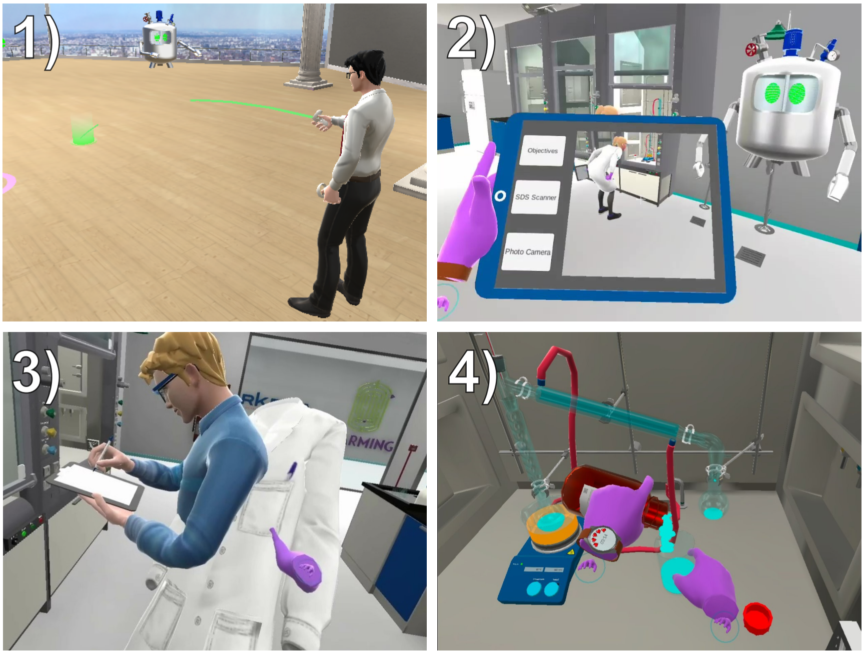


Figure 3.2: In-game screenshots of the VR LaboSafe Game: 1) Tutorial of teleportation; 2) Mission 1: Risk spotting; 3) Mission 2: Risk minimisation; 4) Mission 3: Safe experiment.

3.2.4 Cognitive instructional design

With VR environments and games, there is a high possibility to exceed the cognitive load of the learner when the amount of information is too overwhelming (Kalyuga & Plass, 2009; Makransky, Terkildsen, & Mayer, 2019). The cognitive load is the demand or capacity of cognitive resources that is involved in learning and reasoning (Sweller et al., 1998). There are three types of cognitive load (Sweller et al., 1998): intrinsic, germane and extraneous cognitive load. *Intrinsic cognitive load* refers to the cognitive demand inherent of the task, this means the mental effort that is required to comprehend the content of information. *Germane cognitive load* refers to the cognitive demand that is needed to process information, constructing mental models and developing automation skills. *Extraneous cognitive load* refers to the cognitive demand resulted from irrelevant information that is not related to the learning content. Thus, in order to increase the effectiveness of VR serious games, the instructional design should consider how to make efficient use of the cognitive resources of the learner.

In context of human cognitive processing, three types of cognitive processing occur while learning, according to the cognitive theory of multimedia learning (Mayer, 2014a): essential, generative and extraneous cognitive processing. In order to efficiently manage the cognitive load of the learner, it is suggested to: manage essential processing, reduce extraneous processing and foster generative processing.

Manage essential processing

Essential cognitive processing refers to cognitive processing in the working memory that is needed to mentally select the visual and verbal information from the learning content. Providing information with a mixed modality (i.e. partly visual and partly audio) is more effective in transferring essential information rather than only one modality (i.e. either visual or auditory) (Sweller et al., 1998). Another method is by dividing tasks to learn a complex skill or knowledge into smaller sections and sequencing it from simple to difficult according to the expertise growth of the learner (van Merriënboer & Kester, 2014). In the VR LaboSafe Game, instructional information is divided in different modalities: visual information by animations and colours; and auditory information by Archy talking to the player or by distinct feedback sounds. Moreover, each mission level is sequenced in smaller levels corresponding to one of the learning objectives in which these levels become more difficult.

Reduce extraneous processing

Extraneous cognitive processing refers to cognitive processing that does not support the learning objective. In the case of VR environments and games, excessive extraneous processing is highly probable because high amounts of distracting details are displayed to the learner (Makransky, Terkildsen, & Mayer, 2019). Highlighting elements, that are relevant to the learning material, could draw the attention of the learner towards these elements and away from other distracting elements (van Gog, 2014). In the VR LaboSafe Game, important information is highlighted with a prominent colour or distinct shape. Spatial sound feedback and Archy pointing at objects can also guide the attention of the player in 3D environments. Moreover, after spotting the lab safety risks, the answers given on the tablet quiz are displayed close to the location of the safety risk. Placing related learning content in close spatial proximity, promotes the spatial contiguity of the learner.

Especially for beginners, a new virtual environment with a high amount of decorations and details might be too overwhelming for them. However, in real-life situations, environment are often distracting. For example, Qvist et al. (2015) noticed that the virtual lab in their application looked too organised and clean, while in reality, laboratories are often cluttered and messy. As beginners ar more acquainted with the virtual environment, perhaps more details can be displayed without straining the extraneous processing too much. An alternative method to highlighting, could be the use of foveated rendering to reduce cognitive load. It is a method that displays the detailed virtual environment where the eyes are focused, but blurred vision in the peripheral area of the eye gaze. This helps to reduce visual information input from surrounding area, thus reducing extraneous cognitive load (Romero-Rondón et al., 2018).

Foster generative processing

Generative cognitive processing refers to cognitive processing aimed at comprehension by organizing and integrating the content into knowledge. Several techniques have been researched that provide guidance to the learner to enhance deep learning of the learning content, such as scaffolding the learning content by providing instructional support for novice learners in the beginning, but that fades away as the learner gains more skill and expertise (Pea, 2004). Another technique is by bringing a sufficiently high variability in learning tasks throughout the whole training experience (van Merriënboer & Kester, 2014). In the VR LaboSafe Game, the game levels is designed in such way that hints are provided when the player is struggling. For more experienced players, these hints are not immediately shown, but can be requested when needed. To implement variability, random lab safety risks are spawned at random locations in the virtual laboratory. Each time the player starts a mission level, a different situation is presented.

3.2.5 Motivational game design

Although the novelty and increased sense of presence of VR technology can be inherently motivating, the interactivity of the player with the virtual environment is also very important for sustained engagement (Checa & Bustillo, 2019). By making the learning environment more like a video game, high interactivity and engagement can be ensured. Researchers have suggested that playing games meant for educational purposes leads to greater involvement with the learning experience and motivation to train longer than with traditional teaching methods (Garris et al., 2002). However, other researchers have mentioned that implementing game-elements does not automatically make the training motivating (Hu et al., 2022; Wouters et al., 2013).

To achieve this high level of engagement, the serious game design should support the motivational needs of the player. Game elements, that are based on the self-determination theory (SDT) of Ryan and Deci (Deci & Ryan, 2004), can sustain the intrinsic motivation of the player by supporting the psychological needs of autonomy, competence and relatedness.

Autonomy

The ability to feel in control of one's behaviour and goals is one of the elements of SDT. A flexible game design, that allows players to make their own choices, creates a more meaningful and motivating experience (Nicholson, 2015). Moreover, allowing players to explore and have a sense of control over the environment, sparks their interest and curiosity of the virtual space (Minocha et al., 2017). In the VR LaboSafe Game, safe or dangerous situations can appear depending on the players' decisions. For example, when nitric acid waste beaker is poured in a container of organic solvent waste instead of the inorganic acid waste, then an explosion happens. Also, in order to promote the autonomy, the missions levels are designed in a way that players are free to explore the virtual lab in search for lab safety risks. This exploration design choice resembles the characteristics of an escape game genre.

Competence

Another element of SDT is the feeling of confidence over one's mastery to overcome new challenging tasks effectively. Providing a challenge scaffolding that tailors the level of difficulty to be not too easy nor too hard for the players, can boost their confidence in their abilities (Csikszentmihalyi et al., 1990). This also means that such game design allows a graceful failure of these challenges making it a part of the learning experience to enhance the players' ability to overcome them the next time (Anderson et al., 2018). In the VR LaboSafe Game, this is related to the scaffolding structure of the game levels as mentioned before. Challenges become more difficult the further the player progresses. However, when players do fail at a task causing an accident, the mission level is not immediately stopped, but might reduce the players' health points. This could prevent players to get immediately demotivated every time they do something incorrect. At the end, they will then receive feedback of what they did wrong and how to do better next time. This feedback also shows learning statistics of their learning progress and comparison with other players, so that this could improve their self-efficacy.

Relatedness

The third psychological need involved in the SDT is the feeling of being socially connected with others. Although not all games can afford multiple players, this satisfaction feeling can also be achieved by meaningful interactions with non-player characters (NPCs) in the game (Ryan & Rigby, 2020). Especially with VR technology, a realistic social presence can be simulated. In the VR LaboSafe Game, Archy the batch reactor robot follows the player as a guiding companion. He gives feedback and hints with a friendly voice in order to help the player progress further. Moreover, there are virtual co-workers whom the player will need to keep safe. These characters makes the chemical lab more crowded and realistic instead of an empty environment.

3.2.6 Virtual reality considerations

Modern immersive VR technologies are capable to transport the user to a virtual simulated environment that can strongly resemble reality. Through these technological affordances, the user experiences the sense of presence, which translates to the 'subjective feeling of being there' (Slater & Wilbur, 1997). Immersion and presence have, in some cases, been reported to positively affect motivation and learning outcomes (Lee et al., 2010). The high level of presence is on one hand achieved with highly realistic visual representations, but on the other hand also realistic interactions in a way that the virtual environment behaves like in the real-world. It is an interplay between sensory and interaction fidelity that affects one's feeling of being inside the virtual environment (Mikropoulos & Natsis, 2011; Walsh & Pawlowski, 2002). Especially Head-Mounted Displays (HMDs) are one of the most immersive VR technologies that can provide a high level of visual and interaction realism, while users are visually closed off from the real-world surroundings (Checa & Bustillo, 2019). These devices have been used for educational purposes resulting in positive learning outcomes (Buttussi & Chittaro, 2018). However, users might become disoriented and develop symptoms of feeling nauseous when visual actions inside the device do not match with the actual physical movement of the human body (Davis et al., 2015). Research has been done to search for solutions to prevent or minimise this simulator sickness. Improving the immersion of the user by using adequate hardware and interactive design considerations seems to reduce these symptoms (Kourtesis et al., 2019).

Immersion

While there are different definitions of immersion in literature, one of the definitions is the technical capability of a system where the user perceives a virtual environment through natural sensorimotor contingencies (Slater & Wilbur, 1997). This means that VR HMDs with more advanced technological features can provide a high level of immersion and reduce symptoms of simulator sickness. Some technological characteristics that can affect simulator sickness are: visual performance, spatial audio and motion tracking quality (Kourtesis et al., 2019). The VR LaboSafe Game uses the Meta Quest 2, which provides high-quality performance and comfort by allowing free movement, untethered from a computer.

Interactivity

The term interactivity refers to the interaction between the user and the virtual environment, allowing the user to influence the environment in real-time (Steuer, 1992). VR technology is able to bring a high level of interactivity with natural and intuitive user interactions. This improves the immersion and reduces simulator sickness (Weech et al., 2019). Moreover, allowing users to freely move in the virtual environment by means of teleportation also prevents symptoms of nausea (Cherni et al., 2020). In the VR LaboSafe Game, players are able to intuitively interact with virtual objects, such as grabbing, throwing, pinching, etc. Moreover, they can teleport to different locations in the virtual environment.

3.3 Developing VR LaboSafe Game with game development tools

The Develop phase has the purpose to generate and validate the learning resources (Branch, 2009). From the perspective of a game life cycle, this phase corresponds to the production phase where the game is developed using software tools such as game engines and 3D modelling applications (Dimitriadou et al., 2020; Ramadan & Widyani, 2013). It is also important to decide which hardware equipment will be used for the serious game. Depending on the performance of the selected VR headsets, extra considerations would be needed to optimise the game software in order to avoid game performance issues. Although this phase is primarily managed by game developers and artists, modern easily accessible game development software tools allow anyone to develop games on their own. In this section, we specify which tools were used and which resources were helpful to develop VR LaboSafe Game by self-directed learning.

3.3.1 Software tools

For the development of VR LaboSafe Game, we selected inexpensive development software tools that can be easily acquired by downloading these online. There is a variety in game development tools that are needed for different purposes including 3D modelling software, image editing applications, game engine, programming integrated development environments (IDE) and sound editing software.

In the first stages of the development of VR LaboSafe Game, the necessary game assets are designed and created, more particularly, the 3D environments and 3D objects. These assets were first visualised on paper as design drawings before producing the 3D models using Blender (See Figure 3.3). Blender is a free-of-cost open-source 3D modelling software with a large online community of users (Blender, 2022). Other 3D assets were downloaded for free or at a cost from online platforms for 3D models and asset stores (e.g. Sketchfab (Sketchfab, 2022) and Unity Asset Store (Unity Asset Store, 2022)). GIMP was used as a free-of-cost image editing software in order to create textures for the 3D models and GUI images (GIMP, 2022).

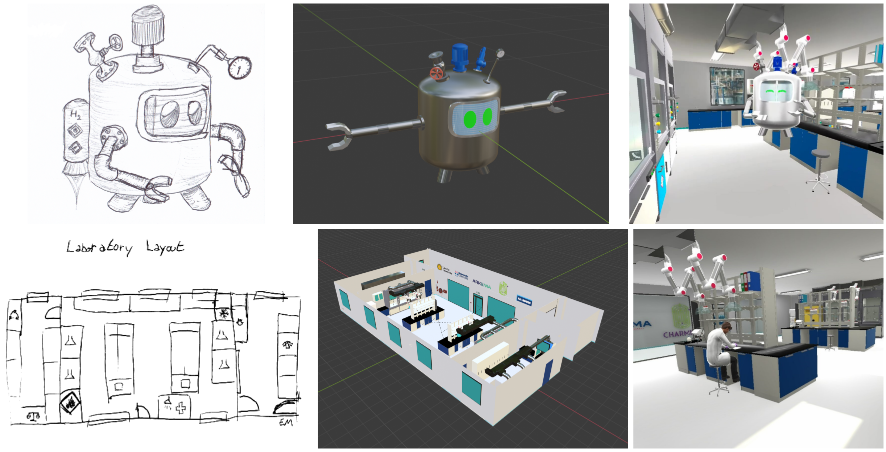


Figure 3.3: The process of (left) designing Archy and the laboratory layout on paper, (middle) 3D modelling this environment and character in Blender, (right) and importing game assets in Unity3D for the VR LaboSafe Game.

Once most assets were obtained, these assets were then imported into the game engine Unity3D for the development of VR LaboSafe Game (See Figure 3.3). Unity3D is a widely used game engine that is based on C# programming language and has a free licensing option (Unity3D, 2022). It offers a lot of advanced options to develop the game with high quality, while also offering an abundance of support. Furthermore, we used Unity's XR Interaction Toolkit that provides ready-to-use solutions for implementing VR interactions in the VR LaboSafe Game more easily.

3.3.2 Hardware equipment

We selected the Meta Quest 2 as VR HMD for the VR LaboSafe Game due to its affordance to move freely while still providing a relatively high performance. This standalone mobile VR headset does not require a cable connection to a computer, allows six degrees of freedom (6DOF) tracking and provides a resolution of 1832x1920 per eye. This device's comfort and high-quality performance could be a way to reduce the symptoms of simulator sickness. However, compared to other high-end VR HMDs that are tethered to a gaming computer, the Meta Quest 2 has hardware components similar to those of a powerful smartphone. This means that this VR headset has a lower processing power than tethered devices, which increases the need for optimisation considerations that will lower the visual quality and visual fidelity of the game. Nevertheless, better comfort and simplicity is preferred over quality in order to make the VR serious game more accustomed to a wide range of population, including people who are not used to VR or gaming in general.

3.3.3 Optimisation techniques

When developing VR applications for standalone VR HMDs, such as the Meta Quest devices, decisions must be made early on in the development stage to optimise the performance of the application. VR games that are not well optimised result in a lagging performance in frames per second. A large frame rate drop could lead to increased simulator sickness (Kourtesis et al., 2019). For the development of VR LaboSafe Game, some common game optimisation techniques are applied (Ferreira, 2019):

- ➢ Using a minimum amount of materials and textures for 3D objects by using texture atlases (i.e. combining multiple textures in a single image file) (See Figure 3.4);
- ➢ Reducing polygon count of 3D models (i.e. minimising the geometric faces that a 3D model is composed of);
- ➢ Using baked lightmaps and avoiding dynamic lighting (i.e. lighting of the environment is fixed and is predetermined in a single image file);
- ➢ Enabling occlusion culling during gameplay (i.e. objects are not rendered when they are not in the player's vision or when they are hidden behind another object) (See Figure 3.4).

3.3.4 Game development by self-directed learning

Game development does not always require a large team of developers. Indie games, for example, are commonly developed by one person or a small team with a versatile skill set. So, developing a VR serious game on your own is an achievable goal, especially with the software tools mentioned in Section 3.3.1. However, a limited expertise and experience of game development can make this goal more difficult to achieve. Fortunately, there is an abundance of free online resources that can support novice developers to self-educate how to develop a game from the beginning. This learning method, where learners take control over their own learning experience, is called self-directed learning (Rashid & Asghar, 2016). Tutorial videos and documentation manuals can be very helpful to master the different game development software tools, while community-driven question-and-answer websites are useful for troubleshooting specific development issues. Table 3.1 shows several online resources that aided the process of developing the VR LaboSafe Game.

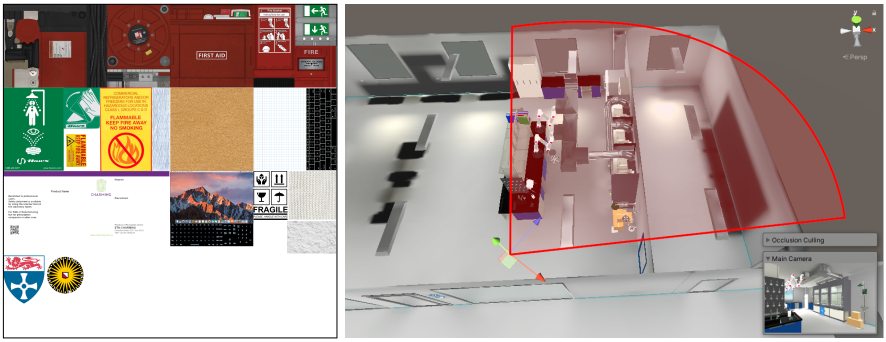


Figure 3.4: Optimisation techniques used in VR LaboSafe Game: (left) texture atlas for a material; (right) occlusion culling of objects behind a wall in the virtual laboratory.

Table 3.1: Online resources used for self-directed learning of development of VR LaboSafe Game.

Software/intention	**Online resources**
Blender	- Youtube channels (e.g. (Blender Guru, 2022))
Unity3D	- Youtube channels (e.g. (Brackeys, 2022), (Code Monkey, 2022), (Unity Youtube, 2022)) - (Unity Documentation, 2022) - (Unity Learn, 2022)
C# programming	- (C# Documentation, 2022)
VR development	- Youtube channels (e.g. (VR with Andrew, 2022), (Valem, 2022))
Troubleshooting	- (Stack Overflow, 2022) - (Unity Answers, 2022)

3.4 Conclusions

VR technology provides the possibility to virtually train hands-on skills in a realistically simulated environment and games are capable of keeping the player constantly engaged. These technological tools could greatly improve current training methods for laboratory safety. However, designing and developing a VR serious game is not an easy process; it is highly multidisciplinary and it is hard to find information in literature on this entire process. Therefore, this chapter presented the whole process with multiple perspectives on how to start from an initial concept to a fully developed VR serious game. The design and development process is shown of VR LaboSafe Game – a VR training tool for chemical laboratory safety – as example, but the process is suitable for any VR serious game.

The game design of VR LaboSafe Game considers an integration of both instructional design and game design. Design principles were taken into account on how to efficiently manage cognitive load of learners and on how to intrinsically motivate players. Furthermore, specific VR design guidelines are presented in this chapter to prevent severe symptoms of simulator sickness due to the use of VR headsets.

The VR LaboSafe Game is then developed using inexpensive and easily accessible software tools, such as Blender and Unity3D. Game performance issues causes lower frame rates, inducing more severe symptoms of simulator sickness. Therefore, game development techniques for performance optimisation, such as texture atlasing and baked lightmaps, are applied to reduce these issues. The development process of the game was heavily supported by tutorials and online resources in order to self-learn game development skills. Design guidelines and development resources mentioned in this chapter can contribute as a guide and as a worked example for future designers and developers of VR serious games.

CHAPTER 4

Usability and Simulator Sickness Testing of using the VR LaboSafe Game

During the development of the VR LaboSafe Game, it is a first priority to ensure that the game is an easy-to-use system and induces minimal symptoms of sickness during the VR experience. A difficult-to-use VR training that causes severe simulator sickness symptoms is not desirable and could negatively affect the results of further evaluation studies. Therefore in this chapter, studies on usability and simulator sickness are described where VR LaboSafe Game is implemented and tested on academic and industrial populations. A first version of the VR game was tested in order to assess whether the usability and simulator sickness were acceptable. Then, based on these results, an updated version of the VR LaboSafe Game was tested in order to verify if the adjustments have improved the usability and simulator sickness of the game.

This chapter is partially adapted from: Chan P., Van Gerven T., Dubois J.-L., Bernaerts K. (2021). Design and Development of a VR Serious Game for Chemical Laboratory Safety. 10th International Conference on Games and Learning Alliance (GALA), Online. Springer International Publishing, Cham. 1: 23-33.
Author's contribution: Chan P. tested the VR LaboSafe Game on academic and industrial populations, and drafted the manuscript. Bernaerts K., Dubois J.-L. and Van Gerven T. managed and supervised the project.

4.1 Introduction

Typically during the production stage of a game life cycle, tests are occasionally performed in order to assess the usability of the developed system on the target population (Arnab & Clarke, 2017; Dimitriadou et al., 2020). According to the International Organization of Standardization 9241-11, usability can be 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." (ISO, 2018). This means that in context of system usability, good usability is achieved when the system lets people complete their goals with reasonable accuracy, with minimal time and effort, and when people are satisfied using the system. From a user perspective, an easy-to-use system is an important aspect of usability. Results of usability tests give more insight to designers and developers what usability issues are present and how to improve the system in future development iterations.

Especially for VR technologies for training purposes, good system usability is crucial because specific learning outcomes are expected by using these systems. Bad system usability can impair the learning processes of trainees by causing overwhelming cognitive load (Hollender et al., 2010). In this case, learners are not only required to invest mental effort to acquire knowledge and skills from the educational software, but also require to learn how to use the VR system. If a system is too complicated and difficult to control, then users might find it unpleasant and are unlikely to use it again (Harris et al., 2020). In particular with VR tools, these devices can cause simulator sickness symptoms, which result in a negative effect on user experience (Somrak et al., 2019). Thus, uncomfortable experiences will reduce the users acceptance for the technology and will negatively impact the motivation and engagement (Mallam et al., 2019).

For these reasons, we must ensure that the usability of the VR LaboSafe Game is optimal and that the severity of simulator sickness is minimal before performing further evaluations. At first, a preliminary version of the game (version 1.0) was tested to evaluate the usability and simulator sickness. Based on these results, adjustments were implemented in the updated version of the game (version 2.0). Then, VR LaboSafe Game version 2.0 was tested to investigate improvements on system usability and simulator sickness and compare the results with version 1.0. Eventually the game version with a better usability and minimal severe simulator sickness symptoms, can then be used for further evaluation studies.

4.2 Testing VR LaboSafe Game version 1.0 (VRLSG-1.0)

The VR LaboSafe Game version 1.0 consists of the first two tutorial levels (i.e. VR interactions and laboratory tablet tutorials) and the first mission level (i.e. Mission One: Risk Spotting). Because animation of virtual characters and voice acting are time-consuming to develop in a game, this first version has only textual step-by-step instructions in the tutorial levels and no animated characters. A first test was then conducted in order to assess if the game has good usability with textual instructions and does not induce severe simulator sickness symptoms.

4.2.1 Methodology and participants

A total of 10 participants (5 women, 5 men, age 20-30) were recruited, who were assigned as interns or students at a research centre of the chemical company Arkema in France. These participants were invited via e-mail to voluntarily participate in this study. Some supervisors of these participants also highly recommended them to participate. Only three participants said they had prior experience with a VR HMD. This test is coded as: VRLSG-1.0.

During the testing sessions, they played the first two tutorial levels and the first mission level of the VR LaboSafe Game version 1.0. The VR LaboSafe Game version 1.0 contained tutorial levels with textual step-by-step instructions and no animated characters. Because of the COVID-19 sanitary measures, the participants wore a face mask and the VR HMD, the Meta Quest 2, was disinfected before each use. The participants played for a duration of approximately 40 minutes continuously without breaks from VR. No one of the participants dropped out before the end.

In order to analyse the usability and simulator sickness, we used the questionnaires System Usability Scale (SUS) (Kennedy et al., 1993) and Simulator Sickness Questionnaire (SSQ) (Brooke, 1995) after playing the game. Additional questions were added on the usefulness of the game, their perceived learning and intention to use. Although most of the participants were French, they stated that they had no issues playing the game and replying to questionnaires in English.

4.2.2 Results and discussion

The SUS questionnaire contains 10-items on a 1 (strongly disagree) to 5 (strongly agree) Likert scale. The calculated SUS score can range from 0 to 100, wherein values of above 68 have an acceptable usability with minor changes required, while systems with values below 68/100 have an unacceptable usability with major changes required (Brooke, 1995). The SSQ contains 16-items on a 0 (none) to 4 (severe) Likert scale. The responses are then calculated to generate the total score for simulator sickness using the equation mentioned by Kennedy et al. (1993). Traditionally, SSQ scores above 20 would be identified as a “bad simulator” (Kennedy et al., 1993). However, a recent meta-analysis mentions that this threshold of 20 SSQ score is outdated and a new evaluation scale needs to be implemented (Caserman et al., 2021). In this literature review, at SSQ scores of 40 or higher, withdrawal rates of approximately one third were observed. For this reason, Caserman et al. (2021) suggest that for modern VR HMDs, it would be more precise to shift this SSQ score threshold to 40 instead (Caserman et al., 2021). This Results of the SUS and SSQ scores are presented per participant in Figure 4.1.

The total SUS scores of the VRLSG-1.0 test vary widely from 40.00 to 72.50 with an overall mean score of 58.25 (SD = 12.8) among the 10 participants. This score is below the usability threshold of 68/100 according to Brooke (1995). The best scoring SUS item states that VR LaboSafe Game is ‘well integrated’, whereas the worst scoring item states that the participants ‘would need the support of a technical person’. The varying SUS scores show that the usability of VR LaboSafe Game should be improved in a way that supports the users better.

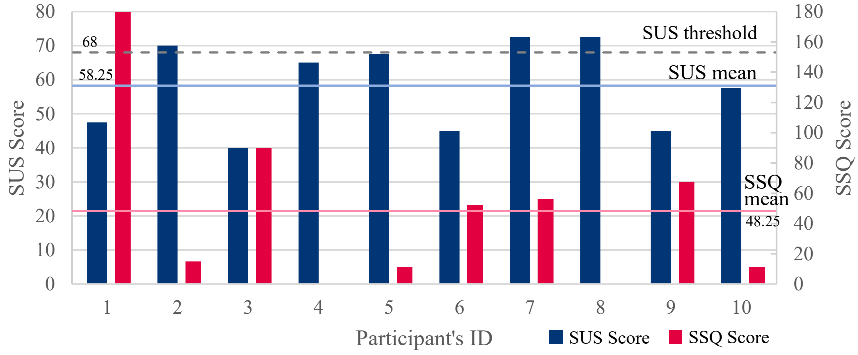


Figure 4.1: System Usability Scale (SUS) (blue) and Simulator Sickness Questionnaire (SSQ) (red) scores per participant.

Indeed, it was observed that some participants had issues with controlling the VR interactions more confidently, especially for people using VR for the first time. Support from the experimenter was often required because the participants did not read the textual instructions or could not fully understand it. This could be caused by the abundance of visual and textual instructions in another language resulting in a high impact on the cognitive load of the user (Low & Sweller, 2014). This usability testing informed that VR LaboSafe Game 1.0 could be improved by replacing most textual information with spoken instructions via a pedagogical agent and add more comprehensible animations demonstrating the VR controls.

In terms of simulator sickness, the VRSLG-1.0 test obtained a mean total SSQ score of 48.25 (SD = 55.68). There are 5 participants who scored an SSQ score below 40, while the other 5 participants scored higher SSQ scores. The most frequently reported (6/10) symptoms are 'eye strain' and 'blurred vision', but also other symptoms, such as 'general discomfort' and 'difficulty focusing' (5/10). Despite no one dropping out and symptoms of simulator sickness were not apparent from our observations, some participants have experienced mild symptoms of visual discomfort according to the SSQ results. This could be explained by the relatively long duration of 40 minutes continuous VR experience. Prolonged duration of visual exposure to a digital screen can cause ocular sickness symptoms, such as eye strain and headaches (Hirzle et al., 2021). Therefore, it is important for the future use of VR LaboSafe Game to allow frequent breaks of a few minutes, especially for first-time VR users in order to minimise discomforting symptoms.

4.3 Testing VR LaboSafe Game version 2.0 (VRLSG-2.0) and comparison with test VRLSG-1.0

The results from the VRLSG-1.0 test show that the VR LaboSafe Game version 1.0 could be improved by reducing textual instructions and allowing breaks from VR between game levels. Therefore, in VR LaboSafe Game version 2.0, textual instructions in the tutorial levels were replaced by spoken step-by-step instructions by Archy and NPC animations were added as examples to demonstrate how to move with VR controllers. In addition, prompt messages were added between game levels informing the player to take a break from VR if needed. Then, another test was performed with VR LaboSafe Game version 2.0 in order to assess whether the system usability and simulator sickness are improved compared to game version 1.0.

4.3.1 Methodology and participants

A total of 35 participants (11 women, 24 men, 91% age 20-30) were recruited, who were students or researchers with a background in chemical engineering or chemistry from the university KU Leuven in Belgium. From the 35 participants, 11 people (31%) have mentioned to have used VR HMDs at least once. This group is coded as: VRSLG-2.0. A schematic representation of the procedure of VRLSG-2.0 is shown in Figure 4.2 with the procedure of VRLSG-1.0 as comparison.

During the testing sessions, they played the first two tutorial levels and the first mission level of the VR LaboSafe Game 2.0. The VRLSG-2.0 group did not play the first VR LaboSafe Game version 1.0. The VR LaboSafe Game version 2.0 contained spoken step-by-step instructions by Archy with NPC animations demonstrating how to move with the VR controllers (See Figure 3.2). Also with this test group, the Meta Quest 2 is used and the same COVID-19 sanitary measures were employed. The participants of this group played for a total duration of approximately 40 minutes. However, in this case, after each tutorial or mission level, which takes roughly 15 minutes, participants were allowed to take a 5 minute break from VR.

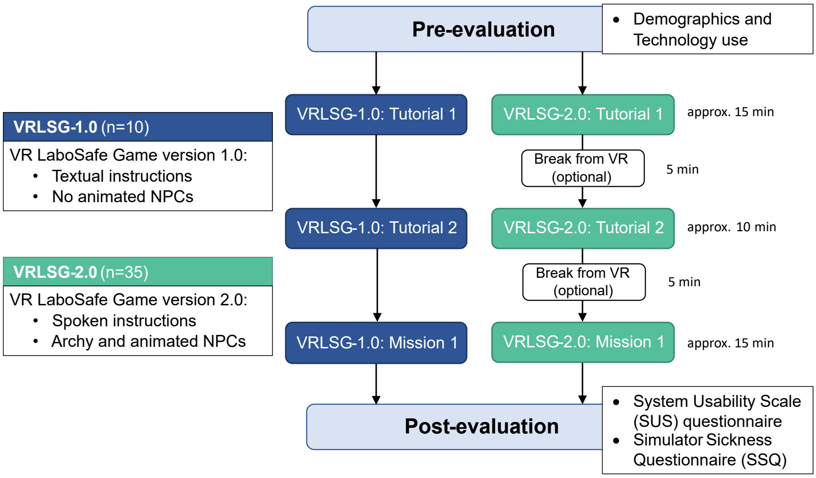


Figure 4.2: Experimental procedure with (left) VRLSG-1.0 group playing VR LaboSafe Game version 1.0 and no breaks from VR, (right) and VRLSG-2.0 group, a separate test group, playing VR LaboSafe Game version 2.0 and with breaks from VR.

In order to analyse the usability and simulator sickness, the same SUS and SSQ questionnaires were used after playing the VR game, like in test VRLSG-1.0. The results of VRLSG-2.0 are then compared with the results from the previous test VRLSG-1.0.

4.3.2 Results

System usability scale

Results and interpretation of the total SUS score of VRLSG-2.0 are shown in Figure 4.3 with the SUS score VRLSG-1.0 as comparison. The total SUS scores of VRLSG-2.0 test vary widely from 37.5 to 85 among the 35 participants with a mean SUS score of 68.5 (SD = 10.9). This score is around the 68/100 usability threshold according to Brooke (1995). Alternatively, Bangor et al. (2009) presented a rating scale to assess the SUS scores with an adjective from ‘worst imaginable’ to ‘best imaginable’. As shown in Figure 4.3, this means that the VR LaboSafe Game 2.0 has an adjective rating between ‘ok’ and ‘good’

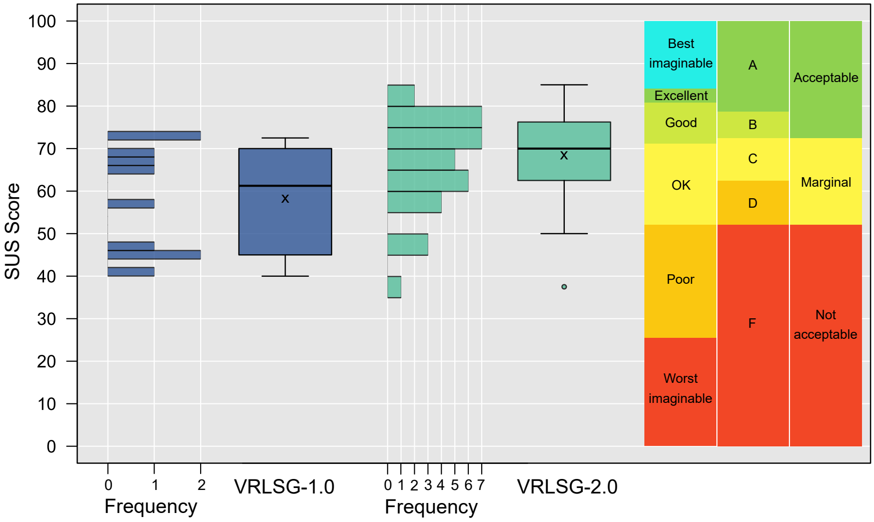


Figure 4.3: Histogram and boxplot of the total SUS score from (left) VRLSG-1.0 group and (right) VRLSG-2.0 group. SUS scores are compared on a rating scale adapted from a SUS analysis toolkit (Blattgerste et al., 2022), which in turn is based the rating scale of Bangor et al. (2009).

system usability and in the range of 'marginally acceptable' to 'acceptable' rating. It is a better rating compared to VR LaboSafe Game 1.0 that has an adjective rating between 'poor' and 'ok' system usability and in the range of 'not acceptable' to 'marginally acceptable' rating.

From Figure 4.4, it can observe that the best scoring items of the VRLSG-2.0 group describe that the VR LaboSafe Game 2.0 is 'well integrated' (S5) and has not 'too much inconsistencies' (S6), whereas the worst scoring items indicate that participants 'would need the support of a technical person' (S4) and that it is rather not completely 'easy to use' (S3).

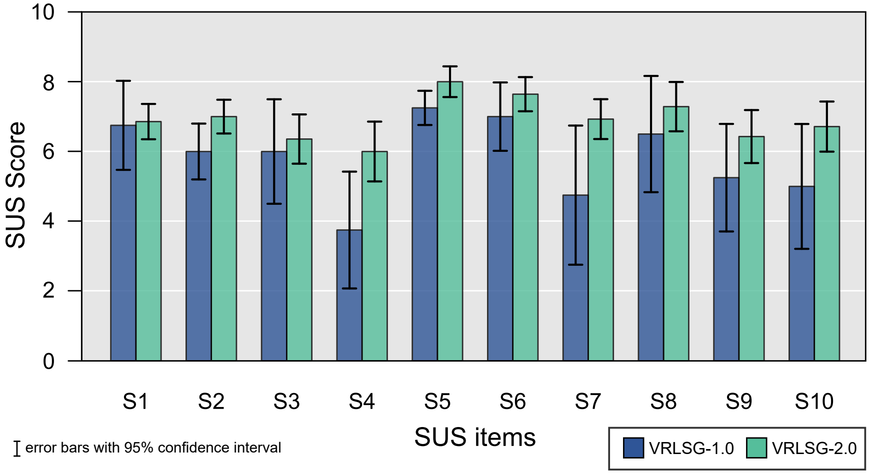


Figure 4.4: Bar plot of SUS scores per questionnaire item from (left) VRLSG-1.0 group and (right) VRLSG-2.0 group, adapted from a SUS analysis toolkit (Blattgerste et al., 2022). These SUS scores are normalised values between 0 to 10, where even numbered items are reverse-coded. SUS item definitions can be found in Appendix B.1.

Simulator sickness questionnaire

Results and interpretation of the total SSQ score of VRLSG-2.0 are shown in Figure 4.5 with the SSQ score VRLSG-1.0 as comparison. The VRLSG-2.0 group obtained a mean total SSQ score of 13.36 (SD = 18.6). There are 17 of the 35 total participants (49%) that reported to have no discomforting symptoms, while 18 of the 35 participants (51%) reported to have some symptoms, which only 4 participants (11%) registered an SSQ score above 40. Among the participants who reported SSQ symptoms, the most frequently reported items are 'general discomfort' (56%) and 'eye strain' (50%), but also 'difficulty focusing' (39%) and 'fatigue' (39%). It should be noted that one person withdrew from the study before finishing the game due to adverse symptoms. It was reported that this person teleported in the environment too quickly causing a sensory mismatch of rapidly changing visual environment and stationary body movement.

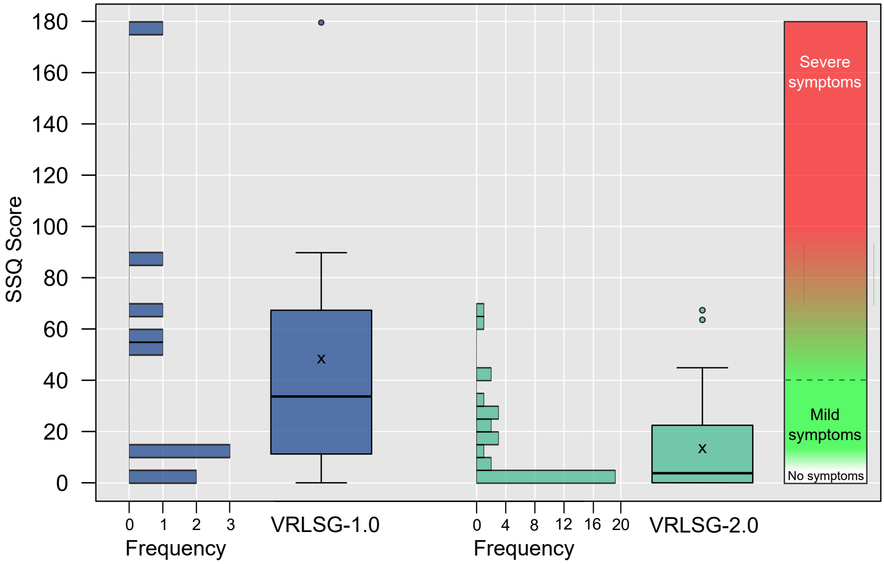


Figure 4.5: Histogram and boxplot of the total SSQ score from (left) VRLSG-1.0 group and (right) VRLSG-2.0 group.

Comparison of textual instruction and no breaks from VR versus spoken instruction + animation with breaks from VR

Because the VRLSG-1.0 and VRLSG-2.0 groups have a large difference in population size, it is necessary to investigate whether both groups have the same variance and are both normally distributed in order to perform parametric tests. Otherwise, non-parametric tests should be considered if these criteria are not met.

With regards to the SUS results, we confirmed that both groups do not have a significantly different variance according to an F-test ($F(9{,}34) = 1.38$; $p = 0.471$) and that both groups seem to have a normal distribution by consulting their histogram, boxplot and q-q plot (See Figure 4.3 and Figure 4.6). Thus, we were able to use a one-sided parametric Student's t-test for unpaired samples assuming that the alternative hypothesis considers that the SUS mean of the VRLSG-2.0 group is larger than the one of the VRLSG-1.0 group. This statistical test showed that the SUS mean of the VRSLG-2 group is indeed significantly higher ($t(43) = 2.52$; $p = 0.008$) than the SUS mean of the VRLSG-1.0 group (See Table 4.1). Furthermore, the calculated Cohen's d effect size shows a large effect size (d = 0.905) in favour for the VR LaboSafe Game version 2.0, hence, there is a good statistical power ($1\text{-}\beta=0.799$) for this comparison. In Figure 4.4, an improvement can be noticed for SUS items such as 'I would need the support of a technical person' (S4) and 'most people would learn to use it very quickly' (S7).

In terms of SSQ results, both groups have a significantly different variance as shown in an F-test ($F(9,34) = 8.97$; $p < 0.001$) and do not have a normal distribution when looking at their histogram and boxplot (See Figure 4.5). Therefore, we used a one-sided non-parametric Mann-Whitney U test for unpaired samples assuming that the alternative hypothesis considers that the SSQ median of the VRLSG-2.0 group is lower than the one of the VRLSG-1.0 group. This statistical test showed that the SSQ median of the VRLSG-2.0 group is indeed significantly lower ($U = 97.5$; $p = 0.0143$) than the median of the VRLSG-1.0 group (See Table 4.1). This means that the group that were allowed to take breaks from VR, showed less symptoms of simulator sickness than the group that did not take a break from VR. Furthermore, a rank biserial correlation (i.e. an effect size measure for Mann-Whitney tests (Kerby, 2014)) of r = -0.44 was found, indicating a medium effect size and a negative correlation between the two groups.

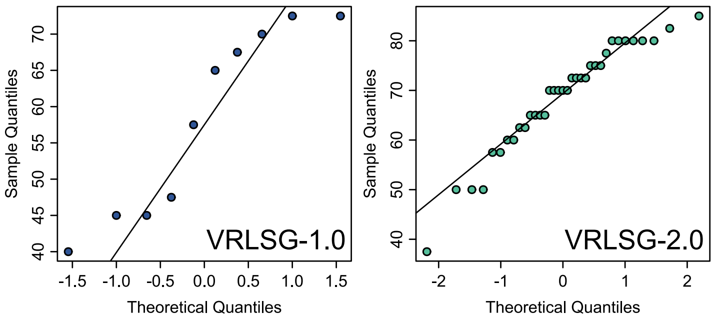


Figure 4.6: Q-Q plots of SUS results from (left) VRLSG-1.0 group and (right) VRLSG-2.0 group.

Table 4.1: Results of statistical tests to compare the SUS and SSQ of group VRLSG-1.0 and group VRLSG-2.0.

Variable	**Mean (SD)**	**Test of equality of variances**		**Test of equal means/median**	
		F(9,34)	Sig.	*t* or *U*	Sig.
SUS (VRLSG-1.0)	58.25 (12.8)	1.38	$p = 0.471$	$t(43) = 2.52$	$p = 0.008$*
SUS (VRLSG-2.0)	68.5 (10.9)				
SSQ (VRLSG-1.0)	48.25 (55.7)	8.97	$p < 0.001$**	$U = 97.5$	$p = 0.014$*
SSQ (VRLSG-2.0)	13.36 (18.6)				

Significance: * p < 0.05, ** p < 0.001.

4.3.3 Discussion

Marginally acceptable system usability of VR LaboSafe Game

It is determined that the VR LaboSafe Game 2.0 has a marginally acceptable system usability with an adjective rating between ‘ok’ and ‘good’ usability. This indicates a significant improvement in usability compared to the VR LaboSafe Game version 1.0 with only textual instructions and without Archy nor other animated characters to guide the users. This result is in line with the modality principle from the cognitive theory of multimedia learning (Low & Sweller, 2014). Users might get cognitively overloaded by the amount of visual information (e.g. text and animation) in the VR LaboSafe Game 1.0. In the VR LaboSafe Game 2.0, this information is distributed to different modalities by reducing the visual information to only animations as well as providing auditory information by spoken instructions from Archy. Consequently, this allows the users to ‘learn to use VR more quickly’ (S7) with version 2.0 than with version 1.0 and would need less ‘support from a technical person’ (S4). However, the overall usability still could be improved more, since items such as ‘need the support of a technical person’ (S4) and ‘it was easy to use’ (S3) have the lowest SUS scores. Perhaps, due to the novelty of VR technology, many users are not familiar with handling such device.

Low simulator sickness effect

When investigating simulator sickness during gameplay of the group who played VR LaboSafe Game 2.0, it is observed that the mean SSQ score is sufficiently low and comparable with other studies using modern VR HMDs (Caserman et al., 2021). Half of the total participants reported no symptoms, while the other half experienced mostly mild symptoms of general discomfort and visual issues. Only 11% of the total participants reported more severe symptoms (SSQ $>$ 40). These reported symptoms are not necessarily associated with feeling nauseous, but could be related to symptoms of digital eye strain or ergonomic issues of wearing a VR HMD (Hirzle et al., 2021). Providing frequent breaks from VR seems to have a significant improvement in minimising simulator sickness symptoms than not taking breaks. Also, while teleportation is the best locomotion technique to reduce simulator sickness (Kourtesis et al., 2019), it is advised not to teleport too quickly after each teleportation, otherwise, adverse symptoms could be induced.

4.4 Conclusions

The results show that the VR LaboSafe Game version 2.0 has an ‘ok’ or ‘good’ system usability and induces minimal simulator sickness symptoms. When compared with the VR LaboSafe Game version 1.0, the VR LaboSafe Game version 2.0 has a significantly improved system usability score and lower simulator sickness. This means that the usability of VR LaboSafe Game is improved by reducing textual information during tutorial and introducing explaining animations with a guiding virtual character. The results also indicate that simulator sickness should be reduced by allowing users to take frequent breaks from VR. The conclusions from these usability tests provided an improved game version and techniques on how to approach future evaluation studies.

There are some limitations of these studies. For example, the results of the comparison between versions of VR LaboSafe Game could be affected due to the small population of the VRLSG-1.0 group compared to the population of VRLSG-2.0. However, the effect sizes and statistical power were observed to be sufficiently high for these test populations. Another limitation could be that the SSQ survey is not the ideal measure for simulator sickness for VR HMDS when it is rather used for general discomfort (Hirzle et al., 2021). However, this questionnaire is heavily used in literature, which makes SSQ scores easily comparable with other studies. Nevertheless, there is an urge for a better questionnaire that includes digital eye strain and ergonomic issues (Bouchard et al., 2021). Furthermore, both SUS and SSQ questionnaires are popular measurement tools, but are generic and cannot pinpoint certain design issues. A survey with items tailored to specific design choices of the VR game is needed in order to improve the game more effectively. An important remark on this study is that there are confounding variables that could effect the results of the study. For example, the language and culture differences between the two different populations could affect the usability of the VR game differently. However, both groups consisted of people from international backgrounds. Another confounding variable could be that the two groups were tested in a different location. A difference in environment parameters could affect the SSQ differently.

CHAPTER 5

Evaluation of Motivation and Engagement using a VR Serious Game

Following the DBR and ADDIE instructional design process, the Evaluate phase comes after Develop phase. It has the purpose to implement and assess the quality of the learning environment. From the perspective of a game life cycle, it can be compared to the post-production phase where tests are conducted with alpha and beta versions of the game. Therefore, in this chapter, an evaluation study is described that investigates how motivated and engaged employees of the chemical company Arkema are for safety training with a conventional (i.e. video lecture) method and with a VR serious game (i.e. VR LaboSafe Game). Moreover, opinions of the employees were gathered in order to obtain a more clear insight with a person-centred perspective on the use of VR serious games as a tool for lab safety training.

This chapter is submitted as: Chan, P.; Van Gerven, T.; Dubois, J.-L.; Bernaerts, K., Study of motivation and engagement for chemical laboratory safety training with VR serious game, Safety Science.
Author's contribution: Chan P. evaluated the VR LaboSafe Game on academic and industrial populations, and drafted the manuscript. Bernaerts K., Dubois J.-L. and Van Gerven T. managed and supervised the project.

5.1 Introduction

Conventionally, safety training is taught with training methods such as classroom lectures, videos and printed safety manuals. These methods include a unidirectional flow of information where the trainee is required to pay attention and listen to the instructor (Bhide et al., 2015). It has the strength that a great amount of theory can be given in a short period and for a large audience (Blair & Seo, 2007). However, trainees are passive in their learning process and this could lead to boredom and reduced attention, which in turn leads to ineffective training (Fivizzani, 2005). Other common safety training methods are on-the-job and hands-on training, where the trainee learns the necessary safety measures by hands-on activities supervised by more experienced workers. This method encourages the trainees to be active in their learning process and cultivates their decision-making skills through experience (Bhide et al., 2015). However, training of highly dangerous situations is not allowed with this method because this puts them and others at a high risk.

VR technology is seen as a simulation-based training method that resembles hands-on training, as the trainees are actively performing safety practices in realistic work environments (Bhide et al., 2015). The trainees are situated in a virtual environment so that hands-on training becomes possible without real-life hazardous consequences. This means that they can make mistakes and learn from these mistakes without jeopardising their own safety, the safety of others and/or jeopardising the integrity of equipment or plant.

In a meta-analysis by Burke et al. (2006), training methods are differentiated based on the participation of the trainee in the training process. Conventional safety training methods (e.g. classroom, video lecture) are categorised as ‘low-engaging methods’, while hands-on training and simulations are classified as ‘most engaging methods’. This study further revealed that the most engaging methods are more effective in reducing negative outcomes, such as accidents, than low-engaging methods. It shows the importance of the level of engagement and involvement in a safety course. However, the classification of engagement between training methods is originated by subjective perspectives from the study’s authors. Mariani et al. (2022) recognise that there is a need to further explore the engagement for safety training from the trainee’s perspective by evaluating the attributes of engagement. Furthermore, there is little evidence how employees are motivated to attend safety training and how the level of engagement of different training methods can play a role in this motivation.

5.2 Theoretical background: safety motivation and engagement

5.2.1 Safety motivation

The general definition of safety motivation is defined as the willingness of an individual to put effort to enact safety behaviours in order to eliminate or reduce the risk of incidents at work (Griffin & Neal, 2000). In the study of Scott et al. (2014), this safety motivation is further distinguished in different levels and types of motivation according to the Self-Determination Theory of Deci & Ryan (2000): intrinsic safety motivation, identified safety regulation, introjected safety regulation, external safety regulation and amotivation. The definitions of these motivation types are presented in Table 5.1. Intrinsic safety motivation and identified safety regulation are further grouped into autonomous safety motivation, while introjected and external safety regulation are grouped into controlled safety motivation. With autonomous safety motivation, employees are self-motivated to work safely because they believe that these activities coincide with their own personal values and interests. On the contrary, controlled safety motivation describes that employees perform safety-related activities because they feel pressured or obliged by their peers (e.g. supervisors, co-workers or organisation). Although safety motivation originally refers to 'working safely' in the study of (Scott et al., 2014), the current study adapts this classification but the 'motivation to attend safety training courses' is investigated instead.

In general, it is preferred to promote the autonomous motivation of employees in order to establish a better safety culture, because this motivation type predicts safety participation (i.e. participating in voluntary activities that support the company's safety culture) (Hedlund et al., 2016; Scott, 2016). Furthermore, safety training was found to be the most important management practice to mediate better safety motivation (Vinodkumar & Bhasi, 2010). Therefore, such training should be designed in a way to stimulate autonomous motivation of the employee.

5.2.2 Engagement during safety training

Engagement of an individual is a complex and broad concept with many different definitions. In literature, there is no clear consensus of the construct because the meaning of engagement can change depending on the object of engagement, the degree of engagement and whether we are talking about engagement during or outside the activity (Ashwin & McVitty, 2015; Bond et al., 2020; Casey et al., 2021). The aim of the current study is not to investigate the full construct of engagement. Hence, specific attributes are selected that could be more suitable for the engagement during safety training. Similarly to the study of Mariani et al. (2022), these selected attributes are based from the User Engagement of O'Brien and Toms (2008) and are related to the three domains of engagement (i.e. cognitive, behavioural and affective engagement) (Ben-Eliyahu et al., 2018): attention, control and interactivity, and reengagement. The definitions of these engagement attributes are presented in Table 5.2. Often, engagement is confused with motivation and used interchangeably. However, motivation is rather the antecedent and driving force for engagement (Bond et al., 2020). Motivation can be considered as an attribute of affective engagement because it contains positive emotions such as, enjoyment and interest. In the remainder of this chapter, when the term 'engagement' is referred to the three attributes mentioned above.

Table 5.1: Definitions for the subscales of safety motivation (Scott et al., 2014).

Safety motivation	**Autonomous motivation**	**Intrinsic safety motivation:** employees engage in safety behaviour completely volitionally because the employee finds pleasure, satisfaction and interest in it.
		Identified safety regulation: employees engage in safety behaviour because they personally believe safety is important for their work environment, not necessarily because they feel they are obliged nor because they have fun doing them.
	Controlled motivation	**Introjected safety regulation:** employees engage in safety behaviour because they feel an internal pressure to behave safely. This feeling can be in the form of guilt or shame.
		External safety regulation: employees feel external pressure or obligation from someone or something else. An external stimulus (i.e. reward for good behaviour or sanction for unsafe behaviour) can motivate them to enact safety behaviour.
	Amotivation	Employees have no motivation to enact in safety behaviour because they feel no reason to do so.

Table 5.2: Definitions for the subscales of engagement.

Engagement	**Attention**	The ability to invest mental effort or focused attention in the safety training. It includes that people are in a state of flow, which is a state when people are so engaged in a task that they devote their total attention in the activity and lose their sense of time (Csikszentmihalyi et al., 1990; Magyaródi et al., 2013). This attribute belongs to the cognitive engagement.
	Control and interactivity	The ability of feeling ‘in charge’ over the activity and the degree of interaction between people and systems. In context of safety training, this also refers to learners making their own instructional decisions resulting in an active involvement in their learning process (Lee et al., 2010). This attribute belongs to the behavioural engagement.
	Reengagement	The degree to which the participant has the intention and desire to do the activity again in the future. It is an important aspect in the process of engagement because, when people are willing to engage with the activity again, this means that they had a positive experience with it or that it offered something new that cannot be obtained somewhere else (Makransky & Lilleholt, 2018; O’Brien & Toms, 2008).

5.3 Study aim and research questions

The scope of the evaluation study in this chapter is to investigate more person-centred variables, namely motivation and engagement. Analysing these variables gives a better perspective of the personal attitude and expectations on safety training given in a conventional way and in a VR method. On one hand, we address: “How can the motivation of employees be described when safety training is given with a more conventional method?”, “How engaged are they during the training process?”. On the other hand, we also address: “How does this motivation change when they play a VR safety training game?”, “Are they more engaged to safety training with the VR game than with a conventional method?”. To answer these questions, the present study makes a comparison of motivation and engagement between a safety video lecture (as an example of a conventional method) and a VR serious game (i.e. VR LaboSafe Game) that is given after the video lecture. The current study will contribute to a better understanding of motivation and engagement of learners for chemical lab safety training.

5.4 Methodology and participants

5.4.1 Participants

The sample population consisted of 37 employees (14 men and 23 women) at a research centre of the chemical company Arkema in France, who were randomly selected and voluntarily agreed to participate. The site director distributed the recruitment invitations to all employees of the site. Some employees participated by their own personal choice, while others were highly recommended by their managers. Most of the employees were either laboratory technicians or managers who have experience in working in a chemical laboratory. The ages of the participants were recorded from 20 to 60 years old and were divided in age groups with an interval of 10 years (See Table 5.3). A total of 14 (38%) participants have responded that they have used VR headsets at least once and a total of 11 (30%) participants have responded that they have played video games before.

Table 5.3: Number of participants with VR and/or game experience in each age group.

Age group in years	20-30	31-40	41-50	51-60
Total participants	12	6	12	7
Have used VR (once or more times)	5	3	4	2
Have played video games	4	1	4	2
Interview participants	4	3	4	3

5.4.2 Procedure

Figure 5.1 summarises the testing procedure of the study. At the start, participants followed a chemical laboratory safety training by means of a video lecture as conventional teaching method. To ensure that it is more similar to a classroom lecture, typical functionalities of video media were restricted, for example, pausing and skipping the video. This video lecture consists of a recorded slideshow presentation with a duration of 12 minutes that is presented by the health and safety manager of the research centre. The content of this safety training video contained the basics of chemical laboratory safety: hazard symbols, hazards classification, ventilation equipment, PPE, chemical storage, etc. This video lecture ensures that every participant received the same baseline of knowledge that is required in order to play the VR LaboSafe Game.

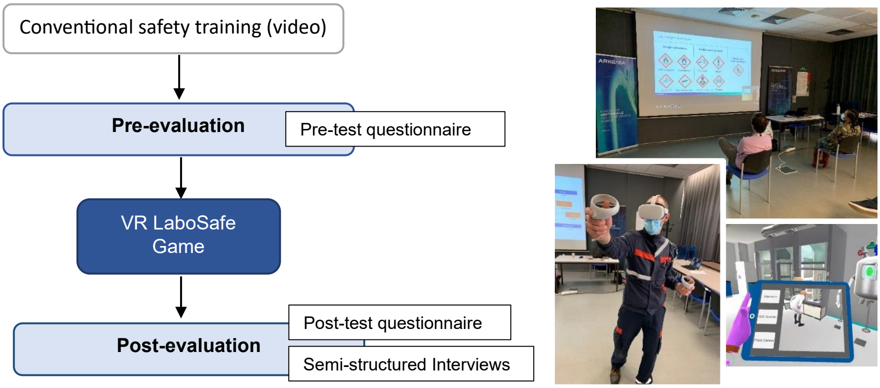


Figure 5.1: Schematic representation of the testing procedure (left) and pictures of the participants (right) following the chemical laboratory safety training via video lecture (top) and VR LaboSafe Game (bottom).

This method is preferred instead of an in-person classroom lecture in order to save time and effort from the instructor. After this video, the participants filled in a pre-test questionnaire about their demographical information, VR and video game experience, attributes of engagement and self-determined motivation of safety training given by a conventional method. The anonymity of each individual was preserved by assigning the participants with a randomly generated ID-number.

Then, the participants play the VR LaboSafe Game that was installed in the Meta Quest 2 VR headset (See Figure 5.1). The total duration of the gameplay was approximately 50 minutes and was divided into separate game levels. The participants played the tutorial levels first and then the first and second mission levels (i.e. 1: Risk Spotting and 2: Risk Minimisation) of the VR LaboSafe Game version 2.0. The game was translated in the mother tongue of the participants (i.e. French). After each level of the game, a participant could choose to take a break from VR in order to prevent severe symptoms of simulator sickness. Because of COVID-19 sanitary measures, the participants wore a face mask and the VR headset was disinfected before each use. After playing the VR serious game, the participants filled in the post-test questionnaire with the same items as in the pre-test questionnaire but now related to safety training with VR serious games. Finally, the participants were then invited for a semi-structured interview to give more in-depth feedback about their engagement and motivation.

5.4.3 Instrumentation

In order to measure the engagement and motivation of the participants for chemical lab safety training, a combination of quantitative and qualitative methods was used. On the one hand, a set of pre-test and post-test questionnaires was used to characterise the motivation and engagement before and after the gameplay. All questionnaire items were scored on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). This Likert scale is a scaling method, so that each item of the questionnaire contains a quantified range of answers. The pre-test questionnaire items were oriented to 'following safety training', while the post-test questionnaire was focussed on 'following safety training with VR serious games'. On the other hand, a semi-structured interview was conducted on one-to-one basis with the participant after answering the post-test questionnaire. This interview was always done by the same interviewer. Questionnaires and interviews were in the mother tongue of the participants (i.e. French). For data processing and reporting, results were translated to English.

Questionnaires on motivation and engagement

The questionnaires were inspired on previously published works. The complete list of items and subscales is given in Appendix C.1 and Appendix C.2.

Engagement during safety training was determined by assessing the three attributes: absorption in the task (i.e. attention); control and active learning (i.e. control and interactivity); and behavioural intention (i.e. reengagement). Absorption in the task was measured by a 4-item scale adapted from the Flow State Questionnaire (Magyaródi et al., 2013). Control and active learning was measured by a 4-item scale adapted from Lee et al. (2010). Behavioural intention was measured with a 4-item scale adapted from Makransky et al. (2018).

The motivation to follow safety training was measured using a 21-item scale adapted from Scott et al. (2016). The subject of each item in the original scale was adapted to correspond to the current subject of attending safety training with a conventional method and with a VR serious game method. Regarding the composite subscales, autonomous motivation scores were achieved by averaging the subscales of intrinsic motivation and identified regulation, while controlled motivation scores were achieved by averaging the subscales of introjected and external regulation as suggested by other researchers (Vansteenkiste et al., 2009). This was justified in our case as the principal component analyses show a clear drop in eigenvalues (pre-test: 2.67, 2.28, 1.39, 1.32; post-test: 2.70, 2.49, 1.31, 1.14) between the second and third component. The first two components combined explained 57% of the variance in the motivation items of the pre-test questionnaire while 63% of the variance in the post-test questionnaire. Moreover, the internal consistency based on Cronbach's alpha is satisfactory for both autonomous (pre-test: 0.90, post-test: 0.90) and controlled motivation (pre-test: 0.82, post-test: 0.93). This Cronbach's alpha is a measure of how items are closely related as a group. Conventionally, values above 0.70 depict an acceptable internal consistency (DeVellis, 2003).

Semi-structured interviews

After the participants have played the VR LaboSafe Game, semi-structured interviews were conducted with 14 participants (7 men and 7 women). In this way, more nuanced data can be collected on the thoughts, behaviours and feelings of the participants about chemical lab safety training and the use of VR serious games. First, participants were asked whether they find the safety training more engaging with the video lecture or with VR LaboSafe Game. Then, they were asked if they believe they have a more autonomous or controlled motivation for safety training in a conventional way and for safety training with VR serious games. For both questions, they were asked to explain their opinion. At last, they were asked for suggestions to improve VR serious games for chemical lab safety training (e.g. content, implementation).

5.4.4 Data analysis methods

For the statistical analysis of the acquired data, we used the programming language R version 4.2.0 as statistical software. To analyse the results of the motivation and engagement questionnaires, responses per subscale were grouped and averaged. These mean values of the pre-test and post-test questionnaires were compared by using t-tests. For the comparison of motivational subscales, two-tailed paired samples t-tests were performed, whereas one-tailed paired samples t-tests were used for subscales of engagement attributes. The reason for this is because the study of motivation has a more exploratory nature, while the engagement for playing VR LaboSafe Game is hypothesised to be higher than with a video lecture. In addition, Pearson correlation was used to evaluate the relationship between motivation, engagement, age, gender, game and VR experience, and time spent in VR game levels.

Motivation is a complex psychological trait that is unique for each individual. Therefore, motivation was further investigated with a more person-centred approach by determining motivational profiles using a two-step cluster analysis method (Van den Broeck et al., 2013). This approach focuses more on the personal motivational characteristics, rather than the average of the whole test group. First, the responses of the motivation questionnaire were subjected to a hierarchical cluster analysis using squared Euclidian distances and Ward's method via the R package 'cluster' (Maechler et al., 2013). This enables us to find the optimal number of clusters and determine the cluster centres. In the second step, the motivational profiles are determined via a k-means clustering analysis while using the previously obtained cluster centres as initial seed points. This combination of hierarchical and iterative clustering methods is recommended in order to fine-tune the preliminary cluster solution (Moran et al., 2012).

In order to analyse and categorise data from the semi-structured interviews, the qualitative data analysis software NVivo version 1.6.1 was used. Interviews were audio recorded and converted to textual transcripts. Then, interesting segments were coded and categorised in themes.

5.5 Results

5.5.1 Safety training engagement

Table 5.4 summarises the descriptive statistics for the measured attributes of engagement. Correlations and internal consistencies of the scales are shown in Table 5.5. The Cronbach's alpha coefficient of the engagement attributes ranges from 0.72 to 0.90 indicating that the internal consistency of the questionnaire is acceptable (i.e. > 0.70) (DeVellis, 2003). The positive correlations between all attributes confirm that they are part of a larger latent variable: the learner engagement during safety training. The mean Likert scores per item are shown in Figure 5.2.

One-tailed paired t-tests indicate that all investigated attributes of engagement are significantly higher after playing the VR game. The "absorption in the task" shows a significant increase with a large effect size ($t(36) = 6.60$, $p < 0.0001$). Participants indicated that the VR game 'engrossed their attention' (E2) more and that the 'time went faster than they have thought' (E4), despite that the duration of the gameplay is much longer than the video lecture (i.e. 50 min. vs. 12 min.). The "control and active learning" shows a significant increase with a small effect size ($t(36) = 1.97$, $p = 0.028$). This means that participants answered that the VR game 'allows them to be more responsive and active in their learning process' (E5) and that it 'promotes self-paced learning' (E7) better than video lecture. The "behavioural intention" shows a significant increase with a medium effect size ($t(36) = 3.21$, $p = 0.001$). The results show that the participants would like to 'participate in more safety trainings with VR serious games' (E11) and 'more frequently' (E10) than safety training with a more conventional method, such as a video lecture.

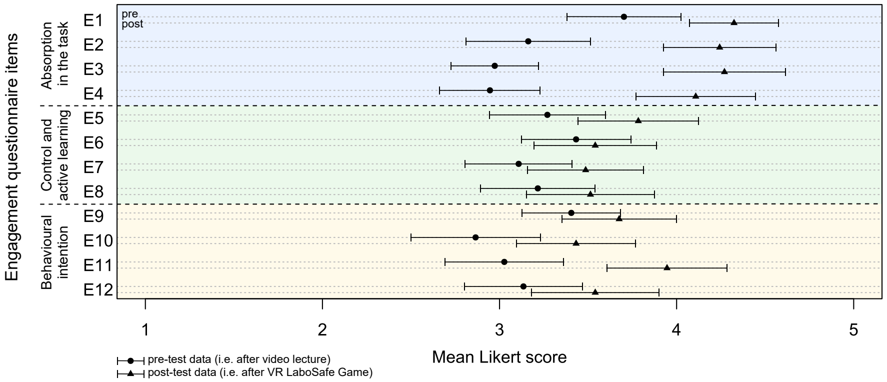


Figure 5.2: Mean Likert scores per item from the engagement questionnaire (circles) after video lecture and (triangles) after VR LaboSafe Game. The error bars denote 95% confidence intervals. The description of the questions (E1 to E12) can be found in Appendix C.1.

Table 5.4: Descriptive statistics and p-values with Cohen's d effect size of the attributes of engagement.

	After video (pre)		After VR game (post)				
	Mean (SD)	Median	Mean (SD)	Median	*t* (36)	*p*	*d*
Absorption in the task	3.20 (0.76)	3.25	4.24 (0.72)	4.50	6.60	<0.0001***	1.08
Control active learning	3.26 (0.79)	3.25	3.58 (0.89)	3.75	1.97	0.028*	0.32
Behavioural intention	3.11 (0.86)	3.25	3.65 (0.93)	3.75	3.21	0.001**	0.53

Note: *t* (36) = t-statistics with 36 degrees of freedom, *p* = p-value of paired sample t-tests, d = Cohen's d effect size.
Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$.

5.5.2 Safety training motivation

Comparison of motivation subscales

Table 5.6 summarises the descriptive statistics for the motivation subscales (i.e. intrinsic motivation, identified regulation, introjected regulation, external regulation and amotivation) and composite subscales (i.e. autonomous motivation and controlled motivation). The Cronbach's alpha coefficient of all motivation subscales ranges from 0.79 to 0.95, indicating that the internal consistency of the questionnaire is satisfactory (DeVellis, 2003). The mean Likert scores per item are shown in Figure 5.3.

When comparing motivation subscales of the pre-test questionnaire (i.e. after video lecture) with the post-test questionnaire (i.e. after VR LaboSafe Game), significant differences can be observed. Intrinsic motivation has a significant increase with moderate effect size. However, also amotivation has increased significantly with a small effect size. Despite this increased value, it should be noted that the Likert scores of amotivation are still in the range of 1 to 2 (i.e. 'strongly disagree' to 'disagree'). These increases are caused by the fact that, compared to conventional training, participants find the VR serious game more 'fun' (M2), while on the other hand, they do not think this VR game is more 'a priority to them (M19) or their workplace (M20)'. When comparing identified, introjected and external regulation, these subscales show a significant decrease with moderate effect size after playing the VR serious game. These significant decreases originate from the fact that, compared to conventional safety training, participants do not think that the VR training game is more 'important for them' (M6), that they will not 'feel more ashamed (M9) or guilty (M10) if they do not follow a safety training with a VR game', and that they are not 'supposed to follow safety training with VR serious games' (M15). When comparing the composite motivation subscales between different training methods, autonomous motivation does not show a significant change, after playing VR LaboSafe game. This can be explained by the increase in intrinsic motivation but a decrease in identified regulation. On the other hand, a significant decrease with a large effect size was observed for controlled motivation, due to the decrease in introjected and external regulation.

Correlations and internal consistencies of these scales are shown in Table 5.5. The results show that positive correlations were found for intrinsic motivation with identified regulation and introjected with external regulation, which confirms that these subscales belong to autonomous and controlled motivation, respectively. As expected, amotivation is observed to be negatively related to autonomous motivation. Furthermore, autonomous motivation subscales (i.e. intrinsic motivation and identified regulation) are moderately to highly positively correlated with all attributes of engagement. Consequently, these attributes of engagement are negatively correlated with amotivation. This means that the higher the engagement during safety training, the higher the autonomous safety motivation and the lower the amotivation.

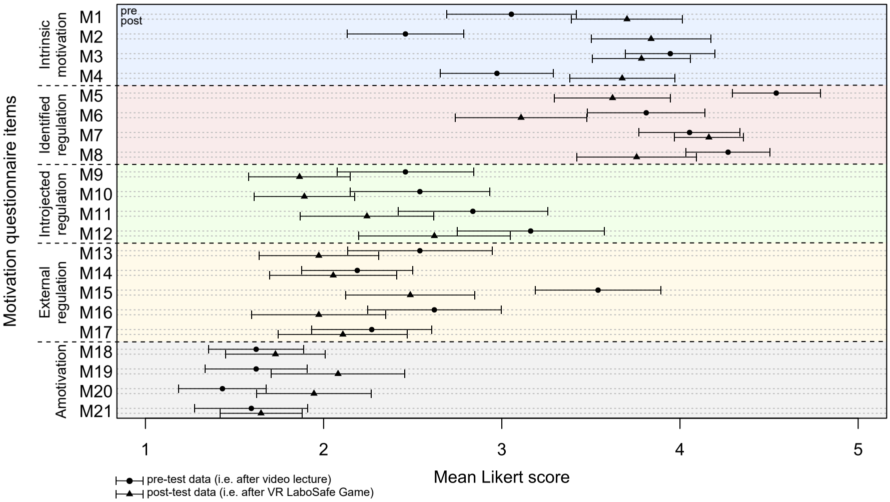


Figure 5.3: Mean Likert scores per item from the motivation questionnaire (circles) after video lecture and (triangles) after VR LaboSafe Game. The error bars denote 95% confidence intervals. The description of the questions (M1 to M21) can be found in Appendix C.2.

Table 5.5: Correlations and internal consistencies among study variables of the pre-test questionnaire (i.e. after video lecture) and post-test questionnaire (i.e. after VR LaboSafe Game).

After video lecture (pre)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Age	-									
2. Gender	**.38**	-								
3. Absorption in the task	-.05	-.13	*(.82)*							
4. Control active learning	-.23	.06	**.63**	*(.82)*						
5. Behavioural intention	-.13	.01	**.72**	**.72**	*(.87)*					
6. Intrinsic motivation	-.30	.02	**.59**	**.60**	**.68**	*(.90)*				
7. Identified regulation	-.22	-.01	**.56**	**.65**	**.62**	**.63**	*(.85)*			
8. Introjected regulation	-.20	.08	.28	.31	**.50**	**.37**	.20	*(.84)*		
9. External regulation	.18	.29	-.18	-.15	-.11	-.09	-.21	**.35**	*(.79)*	
10. Amotivation	.17	.03	**-.54**	**-.33**	**-.52**	**-.52**	**-.73**	-.21	.18	*(.87)*

After VR game (post)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Age	-												
2. Gender	**.38**	-											
3. Absorption in the task	-.08	.12	*(.72)*										
4. Control active learning	**-.39**	-.13	**.62**	*(.85)*									
5. Behavioural intention	-.24	-.05	**.73**	**.77**	*(.90)*								
6. Intrinsic motivation	**-.41**	.04	**.66**	**.78**	**.81**	*(.87)*							
7. Identified regulation	-.26	-.06	**.52**	**.76**	**.74**	**.70**	*(.85)*						
8. Introjected regulation	-.11	-.02	.16	.30	.05	.08	.29	*(.87)*					
9. External motivation	.09	-.03	-.13	.03	-.25	-.31	.03	**.68**	*(.92)*				
10. Amotivation	**.35**	.05	**-.42**	**-.53**	**-.54**	**-.53**	**-.54**	-.03	.21	*(.95)*			
11. VR experience	-.11	-.03	-.08	.15	.15	.29	.13	-.08	-.09	-.07	-		
12. Game experience.	-.01	**.59**	.29	.08	.15	**.34**	.11	-.11	-.13	-.08	.22	-	
13. VR tutorial time	**.42**	-.09	-.10	-.16	-.05	-.24	.10	.27	.03	.02	-.23	**-.34**	-
14. VR mission time	**.45**	-.15	.12	.07	.13	.01	.22	-.03	.08	-.13	-.18	-.16	**.46**

Correlations shown in bold are significant at $p < 0.05$. Cronbach's alpha reliability coefficient for each scale is presented along the diagonal.

Table 5.6: Descriptive statistics and p-values with Cohen's d effect size of the attributes of safety training motivation.

	After video (pre)		After VR game (post)				
	Mean (SD)	Median	Mean (SD)	Median	t (36)	p	d
Autonomous	3.64 (0.71)	3.75	3.71 (0.74)	3.62	0.56	0.579	0.09
Intrinsic	3.11 (0.87)	3.00	3.75 (0.81)	4.00	4.04	0.0002**	0.66
Identified	4.17 (0.71)	4.25	3.66 (0.81)	3.50	-4.15	0.0002**	-0.68
Controlled	2.68 (0.75)	2.70	2.14 (0.87)	2.00	-4.87	<0.0001***	-0.80
Introjected	2.75 (1.03)	2.50	2.16 (0.92)	2.00	-4.74	<0.0001***	-0.78
External	2.63 (0.82)	2.60	2.12 (0.97)	2.00	-3.65	0.0008**	-0.60
Amotivation	1.57 (0.74)	1.25	1.85 (0.88)	2.00	2.25	0.031*	0.28

Note: t (36) = t-statistics with 36 degrees of freedom, p = p-value of paired sample t-tests, d = Cohen's d effect size.

Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$.

Motivational profiles

Clustering analysis of pre-test motivation data (i.e. after video lecture) is performed separately from post-test data (i.e. after VR LaboSafe Game) in order to independently determine the optimal amount of clusters that are distinct from each other. Each cluster displays a unique pattern of scores on autonomous motivation, controlled motivation and amotivation. The distribution of the scores on motivational subscales per cluster can be found in Figure 5.4. In literature, autonomous motivation is considered to deliver more beneficial outcomes than controlled motivation because it allows the satisfaction of the basic psychological needs of an individual (i.e. autonomy, competence and relatedness) (Scott et al., 2014; Van den Broeck et al., 2013). Therefore, clusters are named alphabetically from A to D for pre-test motivation clusters and roman numerically from I to V for post-test motivation clusters in order to depict a ranking from highest to lowest autonomous motivation. When autonomous motivation scores are similar, then the cluster with a lower controlled motivation is more desired because controlled motivation does not satisfy or even inhibits the basic psychological needs (Ryan & Deci, 2000; Van den Broeck et al., 2013).

Concerning the pre-test motivation, the two-step clustering analysis resulted in a four-cluster solution. Inspection of the dendrogram (See Appendix C.1 and the comparison with other cluster solutions indicate that the four-cluster solution is the most suitable. Profile A (n = 8) shows high scores in both autonomous and controlled motivation and can be identified as a 'highly motivated' profile. Profile B (n = 19) shows moderately high autonomous, but low controlled motivation values. The moderately high values of autonomous motivation are mainly attributed to the high identified regulation, not so much to intrinsic motivation. This profile is identified as 'moderately autonomous' motivation profile. Profile C (n = 11) shows neutral scores for autonomous and controlled motivation. This profile has a high identified regulation, but a low intrinsic motivation that is negatively oriented (< 3 Likert score), thus resulting in a neutral autonomous motivation. So, profile C is identified as 'neutral' motivation profile. Profile D only has one participant who reported a very low level of autonomous motivation and neutral controlled motivation. It is noteworthy that this profile exhibits a high amotivation whereas the amotivation in the other profiles is much lower. Hence, this profile D can be identified as an 'amotivated' profile.

Concerning the post-test motivation, the same clustering analysis procedure was repeated which resulted in a five-cluster solution. In this case, the highest ranked profile I (n = 8) is distinct from the highest ranked profile from pre-test motivation (i.e. profile A). People with profile I have a high level of autonomous motivation and a low level of controlled motivation. This profile is similar to profile B (i.e. moderately high autonomous, low controlled) but with a higher quantity. Thus, this profile can be described as 'highly autonomous' motivation profile. Some profiles such as profile II (n = 3) and profile III (n = 13) are not much different from profile A and B from pre-test motivation, except a slight increase of intrinsic motivation and decrease in identified regulation in profile III. Therefore, these remain as 'highly motivated' profile and 'autonomous' motivation profile, respectively. Other profiles such as profile IV (n = 9) and profile V (n = 4) appear to have a different motivation distribution than the pre-test ones. Profile IV shows a much lower controlled motivation, while intrinsic motivation increased and identified regulation decreased, resulting in a more equalised autonomous motivation. This profile can be described as a 'low-controlled' motivation profile. Also, the lowest-ranked profile V shows

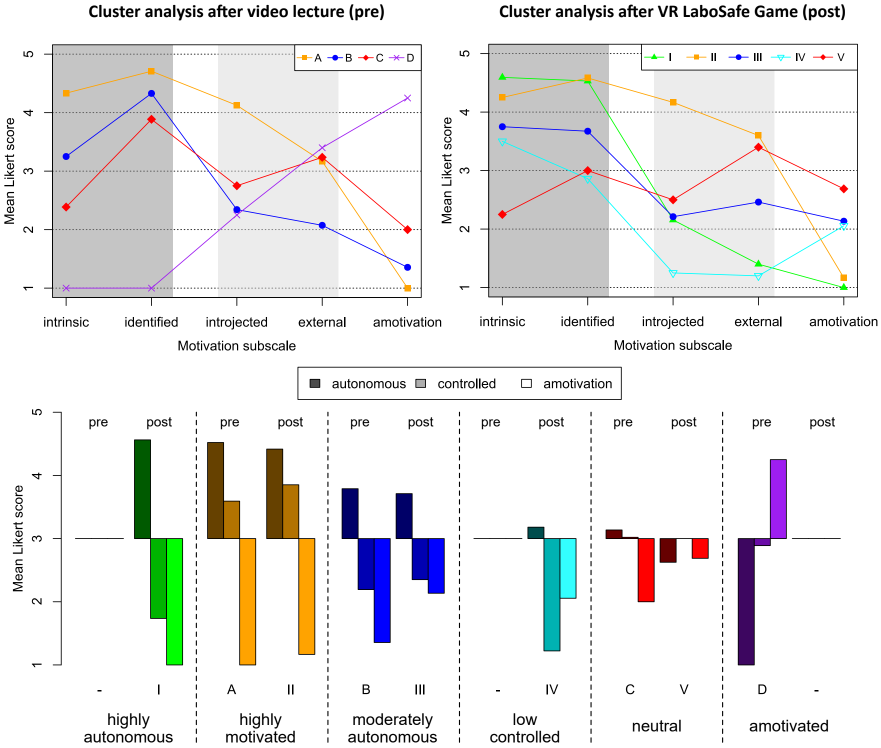


Figure 5.4: Motivational profiles obtained by k-means clustering analysis of motivational subscales in pre-test (i.e. after video lecture) and post-test (i.e. after VR LaboSafe Game). The line curves do not represent continuous data but are intended to guide the eyes.

resemblance to a ‘neutral’ motivation profile with a slightly negatively scored autonomous motivation, while no participants are identified with an ‘amotivated’ profile anymore.

Differences in demographics and engagement attributes between motivational profiles are displayed in Table 5.7. The changes of motivational profile of an individual are visualised in Figure 5.5. After playing the VR serious game, participants from the age group 20-30 years old are identified with a higher ranked motivation profiles (i.e. I, II, III), while 51-60 years old participants are identified with lower ranked profiles (i.e. IV, V). It seems that some of the participants of the age group 51-60 years old have moved from high ranked profiles (i.e. A, B) before playing the VR serious game to low ranked profiles (i.e. IV, V) after playing the VR serious game. This is aligned with the results that the age of the participants is negatively related to intrinsic safety training motivation and positively related with amotivation after playing the VR serious game as seen in Table 5.5. Another noticeable difference between the motivational profiles in Table 5.7 is that higher-ranked profiles have higher scores on attributes of engagement than lower-ranked profiles. This coincides with the positive correlation between engagement attributes with autonomous motivation subscales as seen in Table 5.5.

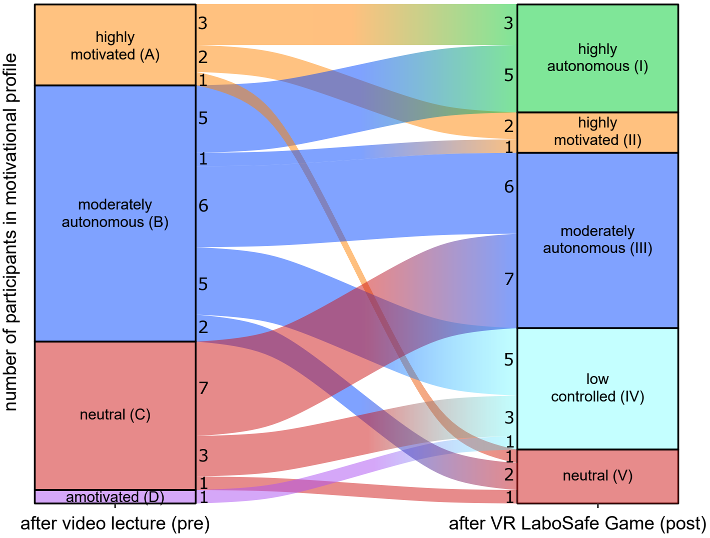


Figure 5.5: Changes in motivational profiles of participants between the video lecture training (left) and the VR-based training (right).

Table 5.7: Differences in demographics and engagement attributes per motivational profile.

	After video lecture (pre)				After VR LaboSafe Game (post)				
	A	B	C	D	I	II	III	IV	V
n	6	19	11	1	8	3	13	9	4
Gender									
Women	3	14	5	1	6	2	8	4	3
Men	3	5	6	0	2	1	5	5	1
Age									
20-30 years old	4	5	3	0	5	2	4	1	0
31-40 years old	0	6	0	0	1	0	1	3	1
41-50 years old	0	6	5	1	2	0	7	3	0
51-60 years old	2	2	3	0	0	1	1	2	3
VR experience									
No	5	9	8	1	4	2	7	7	3
Yes	1	10	3	0	4	1	6	2	1
Game experience									
No	4	14	7	1	5	2	10	5	4
Yes	2	5	4	0	3	1	3	4	0
Engagement attributes (mean Likert scores)									
Absorption in the task	4.08	3.25	2.75	1.75	4.75	4.97	4.00	4.19	3.56
Control and active learning	3.92	3.28	3.07	1.00	4.28	4.50	3.67	3.00	2.50
Behavioural intention	4.17	3.13	2.61	1.75	4.59	4.08	3.58	3.25	2.56

5.5.3 Semi-structured interviews

Figure 5.6 presents which training method the participants find the most engaging and whether their motivation is more autonomous or controlled to follow safety training with a conventional method or with VR serious games. The opinion of the participants about the teaching methods for chemical lab safety training is then categorised in different themes such as, the strengths and weaknesses of both learning methods and suggestions to improve VR safety training (See Table 5.8).

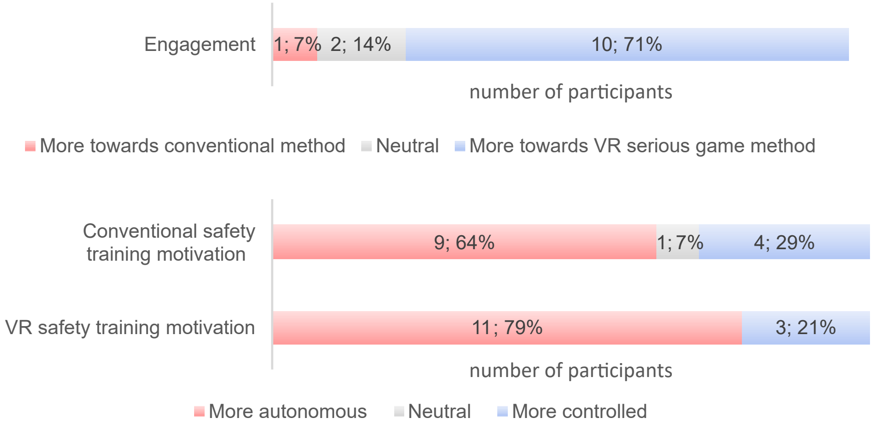


Figure 5.6: Participants' perception of the engagement and motivation of safety training with VR serious games compared to training with more conventional methods.

Table 5.8: Opinions of the participants (N = 14) derived from the semi-structured interviews on the use of VR serious games for chemical lab safety training.

Strengths of safety training with a conventional teaching method	**No. of people**	**Weaknesses of safety training with a conventional teaching method**	**No. of people**
Better knowledge acquisition	6 (43%)	Repetitive content, not easy to adapt	8 (57%)
People already have the habit	2 (14%)	Can lose attention quickly	6 (14%)
Real reality	1 (7%)	Passive learning	3 (21%)

Strengths of VR safety training	**No. of people**	**Weaknesses of VR safety training**	**No. of people**
Situated learning variability	12 (86%)	Novelty usability issues	9 (64%)
Control active learning	10 (71%)	Ergonomic issues	7 (50%)
Absorption in the task	8 (57%)	Less suited for knowledge acquisition	5 (36%)
Novelty	6 (43%)	Discrepancy of reality	3 (21%)
Presence fidelity	4 (29%)		

Suggestions for improvement	**No. of people**
More frequent, smaller sessions	7 (50%)
Complementary (i.e. combination with conventional methods)	7 (50%)
More specialised to the job	4 (29%)

Conventional safety training: engagement and motivation

The interviews with employees of the chemical company showed that, in general, employees are more autonomously motivated to follow safety training courses. Participants have described that safety training is “important for their own safety and safety of others” (P1 - the code refers to a quote from a participant) and that they “want to learn new things and enrich their knowledge about safety” (P8). Some employees are “super voluntary for safety activities” (P9). One interviewee said:

> P6: “In the field of chemistry and for our activities, it’s important to work safely; and that I find it interesting as well.”

Despite the fact that there is an obligation from the company to follow safety training, they are still self-motivated:

> P4: “I am motivated by myself and enjoy doing it, but we still have an obligation to follow, as we work in the laboratory and it is mandatory to be aware of safety.”

However, even though these employees are autonomously motivated, they do not find safety training courses amusing. For example, one employee explained:

> P5: “It’s not amusing, it’s not about pleasure. It is more about being motivated for my own safety and the safety of the people around me.”

This displeasure can be more pronounced with other employees who have a more controlled motivation for safety training:

> P14: “I don’t go too much for my own pleasure. I go there because we have to go there.”

The modest dislike for current safety training arises from the fact that employees believe they receive “too much safety training” (P10) with “repetitive learning content” (P6). Particularly with conventional teaching methods, “content cannot be easily adapted to their work” (P5) and learners are “passive in front of a screen or a tutor” (P1). Consequently, it is “easy to not pay attention” (P7), hence a low engagement during safety training.

Strengths of VR safety training

When participants were asked about how safety training with VR serious games is compared to conventional methods, they mostly responded (71%) in favour of the use of VR serious games in terms of engagement. They mentioned several advantages that VR technology can provide which conventional methods lack.

Firstly, training with VR can provide opportunities for situated learning and variability. Employees described that this technology is “capable of realising almost real situations” (P6) with “more concrete examples” (P2). Compared to conventional methods, VR can “more easily add new things and create different situations each time” (P10), even “dangerous situations without placing the person in danger like in a real laboratory” (P3). Participants believe this training tool makes it possible for them “to visualise better” (P1) which makes it also “better for practising situations” (P4).

> P13: “When you are in the video game, you realise that you are in a realistic situation. When in real-life, you cannot see everything and then you can miss problematic situations.”

Secondly, VR serious games allow learners to have control and be active in their own learning process. Participants mentioned that training with VR serious games is “more dynamic” (P6) and has “more interactions” (P13) which leads to learners being “more active” (P1) and makes them “more an actor of their own training” (P14).

> P3: “You are not sitting and listening, you have things to do.”

This high level of interactivity and activity makes learners “more attentive in the video game, thus it permits to hold their attention better” (P9) which makes “the time pass more quickly” (P14).

> P9: “You keep moving, you keep concentrating. You do not lose yourself in your own thoughts and you are always concentrating on the subject.”

The novelty of VR technology also brings an increased interest than conventional methods. Before the safety training with VR, participants were “curious to see how it works” (P2) and they wanted to “discover another environment” (P14). While playing the VR serious game, participants were surprised by “how advanced this technology has grown” (P1). They felt a heightened presence due to the high fidelity of the tool. It is “much more immersive” (P1) because “the virtual space is unlimited” (P12) in which you can “walk everywhere” (P5) and because “certain situations can come close to reality” (P6).

> P12: “The serious game really shows laboratory safety and that I find very striking. It is really a virtual reality, that’s for sure.”

These strengths of VR safety training improve the autonomous motivation compared to conventional methods because employees believe training with VR serious games is “more amusing” (P5) and “more attractive” (P1). For these reasons, they find the VR safety training more engaging than conventional safety training.

Weaknesses of VR safety training

The participants have also mentioned limitations of using VR serious games for laboratory safety training. While novelty is one of the strengths of VR safety training, it is also its limitation for some users. For example, when asked if their interest would remain after the initial encounter, participants mentioned that “this phenomenon of discovery would fade” (P14). Though, it can still remain interesting when new things are provided each time.

> P10: “In the beginning I was happy to test the VR game, but after, it would become repetitive. That is why I hope VR can bring different things each time. It can create anything you want, so it is more easily to add new things and create different situations each time.”

Moreover, some employees, who have never used VR or played games before, might struggle with the usability of this novel tool. For example, some participants mentioned that “the controllers are not easy to manipulate” (P13) and that they “need to put a lot of effort to learn to play the game” (P8). When trainees do not have the habit to use this new technology, it can become too complicated, making them “frustrated because they do not succeed in doing the things that they wanted to do” (P9). One participant mentioned that it also occurs with older people:

> P2: “I have talked to other people around my age of 50 years old and all of them had the same problems with VR. It is because we

> don't have the habit and it becomes complicated very quickly at some point."

This usability frustration together with wearing "a too heavy VR headset" (P8) can cause slight ergonomic issues. Due to the intense concentration and a lot of movement, participants experienced "tiredness" (P2), "a bit of nausea" (P4) and "reduced spatial awareness" (P3).

> P1: "Sometimes I felt nauseous, but because the headset was quite heavy. I do not have the habit (of using VR headsets), I think."

Another limitation that the participants have noticed, is that using VR serious games is less suited for knowledge acquisition compared to conventional methods. Some responses are:

> P14: "Difficulties of interacting with the game can obstruct taking in information. You can get distracted by other things, thus not learning the safety content."

> P5: "I do not find all the information in the VR game that are taught in a classical lecture."

Moreover, some participants prefer training in real-life environments rather than a virtual reality. For example, one participant described:

> P8: "The virtual reality is an artificial reality that is not the real reality. So, the construction of this reality of the game depends on the constructor and not from reality."

Ultimately, these weaknesses of VR safety training can result in a more controlled motivation and reduced autonomous motivation, as frustration makes the training unpleasant and becomes an obligation to follow.

Suggestions for improvement

In order to overcome the weaknesses of VR safety training, participants were asked to provide suggestions to improve the VR learning experience. Firstly, to overcome the most common issues of using VR (i.e. usability and ergonomic issues), the participants suggest “to give multiple smaller sessions and progressively introduce VR” (P6) and “provide more time to learn VR for those who need it” (P10). Most participants agree that they would get used to using the novel technology after the first sessions.

> P6: “It’s a matter of habit. It is like using a new tool or a new machine; you have to take the time to understand it.”

Secondly, employees still see the value of both conventional and VR teaching methods. They consider these methods as “complementary in a way that one cannot do without the other” (P10). A few suggestions to combine both teaching methods are:

> P1: “Give a small part as a lecture in advance, to teach people what needs to be taught. Then afterwards, the video game can implement small exercises.”

> P9: “A presentation in class before the game and one after as debriefing. This is to make it more interactive between the learners for them to discuss and to make exchanges.”

Other participants suggest to design VR applications with “modules that are more specialised and more oriented to their job” (P9).

> P13: “The video game would be great for safety training of specific products, such as hydrogen fluoride or peroxides, because these products are a bit special.”

5.6 Discussion

5.6.1 Conventional lab safety training: Engagement and training motivation

From the results of this study, employees have generally a high autonomous motivation for conventional safety training. Especially, a high level of identified regulation was observed because employees believe safety training is important for them in order to acquire skills and knowledge to improve their own safety and the safety of others. This is not surprising as the chemical company has a high focus on safety culture (Arkema, 2022a). Even though safety training is mandatory, some employees believe that their motivation is more autonomous rather than controlled. However, other employees can feel more obliged when they do not like the safety training, so motivation subscales can vary from person to person. For this reason, individuals were grouped with similar safety training motivation, resulting in four characteristic motivational profiles. Similar to the study of Howard et al. (2016) and their “motivational profiles at work”, we obtained profiles that are highly motivated, autonomously motivated, neutral and amotivated. These profiles are rated from high to low autonomous motivation, respectively. Interestingly, intrinsic motivation is lower than identified regulation in each profile and amotivation increases with lower-ranked profiles. This can be explained by the fact that some employees do not find conventional safety training pleasing (when it is given with a classroom or video lecture based method). They also believe that they are too passive and that such training is too repetitive and difficult to adapt, hence, they can lose their attention more easily. As autonomous motivation is correlated to the attributes of engagement, the low-engagement characteristic of conventional teaching methods can result in a lower quality of motivation.

5.6.2 Safety training with VR serious games: engagement and motivation

Positive effect of using VR serious games as safety training tool

Engagement and motivation were assessed after the employees played VR LaboSafe Game as training tool in order to compare the lecture-based training method with the VR serious game method. Firstly, the results show that the lab safety training with VR serious game has significantly higher scores on intrinsic motivation and on all measured attributes of engagement (i.e. absorption in the task; control and active learning; and behavioural intention). The interviews with the participants confirm that most of them are in favour of VR serious games than lecture-based methods in terms of engagement. In general, with VR-based training, employees believe that they are more active, attentive and have more control of their learning process. Furthermore, coinciding with the framework of Casey et al. (2021), trainees enjoy that VR has the possibility to realise situations that are more relevant for them and that they are able to be immersed in a semi-realistic environment. However, the novelty effect of the technology plays a large role in this increased engagement which could wear off after multiple uses (Makransky & Petersen, 2021). Nevertheless, employees will still be interested when VR keeps bringing new content every time. So, results of this study concur with the engagement classifications of previous reported literature where lecture-based methods were assumed to be low-engaging methods and VR-type methods as high-engaging methods (Burke et al., 2006; Casey et al., 2021).

When looking at motivation for VR safety training on an individual level, we determined five motivational profiles: highly autonomous, highly motivated, autonomous, low-controlled and neutral. Compared to motivational profiles for conventional safety training, the highly motivational and low-controlled motivational profiles are two profiles that were not described before while no people were identified with amotivated profile anymore. Moreover, with VR serious game as training method, more people (n = 11) are identified with high-ranked profiles (i.e. highly autonomous, highly motivated) than when a lecture-based training method (n = 6) is given. These changes can be related to the aforementioned strengths of using VR as training tool, since these employees find training with a VR serious game more engaging and more amusing while feeling less obliged.

Limitations of using VR serious games as safety training tool

It was observed that motivation to attend a VR safety training is related with age. Younger employees (20-30 years old) have high-ranked motivational profiles, namely highly autonomous, highly motivated, moderately autonomous (i.e. I, II, III). Older employees (51-60 years old) have a lower-ranked motivational profile, namely low-controlled, neutral (i.e. IV, V). Furthermore, age is found to be negatively correlated with intrinsic motivation and positively correlated with amotivation. Based on the interview results and that the age of employees is positively correlated with the time spent in VR, a possible explanation could be that the usability of VR technology is more complicated for older people. This goes in-line with the definitions of 'digital native' and 'digital immigrants'. Digital natives are people, commonly from younger generations in developed countries, who are more proficient with the use of digital technologies. Whereas digital immigrants, commonly people from older generations, are not used to these technologies (Prensky, 2001b). Digital natives have more experience with using the internet, using smart devices and playing video games, thus being more skilful than digital immigrants (Akçayır et al., 2016). This is also reflected in the negative correlation between game experience and the time spent in VR LaboSafe Game tutorial levels. The novelty of VR technology can also bring mild ergonomic issues when users are not used to it. First-time users mentioned that long periods of constant concentration and movement can make them tired, slightly nauseous or lose spatial awareness. Despite these issues of using a novel tool, most employees are still willing to familiarise themselves with VR technologies. They suggested to provide more frequent and smaller session with VR while also introducing the technology gradually for first-time users.

An important observation for chemical companies is that employees do not believe that VR safety training can replace conventional safety training. While a VR serious game is good to practise real-life dangerous situations, it is less suited to learn factual knowledge about safety (Makransky, Borre-Gude, & Mayer, 2019). This attitude about VR safety training could explain the significant lower scores for identified regulation and higher scores for amotivation compared to conventional methods. Employees are used to attend safety training with conventional methods, but are unfamiliar with using VR games as training tool. They are certain that conventional safety training has a high priority for them. However, based on the questionnaire and interview results, it could be that they do not know if VR serious games will be as important for them, since some people have minor issues with the technology. Nevertheless, they suggest VR serious games should be utilised as a complementary tool with lecture-based training. Conventional safety training can be given before the VR serious game to teach theoretical content and can also be given after the VR experience as a debriefing session (Crookall, 2010).

5.6.3 Other aspects of VR serious games as training tool

The term 'conventional training method' is used in this dissertation to depict how safety training has been taught conventionally, namely classroom and lecture-based training given by an instructor. However, currently, there is a trend to support more learner-centred teaching methods where the focus is more on the learner rather than on the instructor's teaching. Without using a VR serious game, there are other ways to add engaging and motivating elements to conventional teaching: adding interactive questions; allowing collaborations between trainees; giving helpful feedback; etc. (Burke et al., 2006; Alaimo et al., 2010). With these techniques there is a possibility to significantly increase the engagement and motivation of the trainees for conventional training methods.

However, a VR serious game as training tool can bring other benefits. For example, for a very large training group, the aforementioned techniques can become more challenging for the instructor to implement while providing every trainee with optimal support. While with a VR training tool, training could be given with better individual support taking into account the available devices and group distributions. Moreover, trainees can follow more training sessions more frequently without requiring more time and effort from the instructors.

Such VR game training also has the advantage to simulate real work procedures with a high level of realism without endangering the trainee. When the work environment and procedural handling are accurately simulated, trainees could practise procedures multiple times and stimulate their spatial and procedural memory of their work (Buttussi & Chittaro, 2018). With these realistic simulations, it would be possible to detect procedural issues and latent design failures in their real work activities, making real procedures more efficient when corrected. However, with current VR technology, there are limitations that restricts the system to be an exact copy of the real world. For example, interactions within the virtual environment are designed and programmed by its creators. This means that unforeseen events, such as unpredicted chemical reactions, would not appear in the virtual world. Furthermore, there is a limit on how many possible outcomes one can program in a virtual simulation (Amokrane & Lourdeaux, 2009). Another limitation is that an experience of the full human sensory input remains very difficult to replicate. With current VR technologies, users cannot touch or smell virtual objects like in the real world (Wu et al., 2019). This is particularly important to detect hazardous chemical odours. Perhaps the implementation of other rapidly evolving technologies, such as AI, haptic/olfactory devices and cloud computing, can mediate the exact replication of the real world in the virtual universe.

Regarding the cost of implementing a VR serious game as safety training tool, it can be very expensive depending on how complex the training should be. The hardware and physical space needed would require a large investment from the companies for such VR implementations. Moreover, additional cost and time is needed for the development of the training software, depending on the development team and the complexity of the software. Therefore, the initial cost for a company to implement VR safety training would be significantly larger than just providing a conventional training with learner-centred teaching methods. However, in the long term period, VR safety training could bring a better return of investments as these applications can be frequently used after initial acquisition, while classroom training needs a constant fee for hiring instructors. This also avoids the effort of recruiting suitable instructors, thus allowing for consistent high quality training even across international departments.

Eventually, in order to consider the implementation of a VR serious game as safety training tool, one should consider: the type and complexity of the training; initial financial investments; frequency and period of usage; and the digital/VR literacy of the trainees.

5.7 Conclusions

This evaluation study shows that lab technicians and managers at a chemical company have a high autonomous motivation to follow safety training in general, particularly the subscale identified regulation. They find it important to acquire the necessary skills to keep a safe work environment for them and their peers. However, with conventional training methods, such as classroom and video lecture, some employees do not find it entertaining to follow because they are rather passive, find it too repetitive and can lose their attention quickly. Results of this study show that when safety training is given with a VR serious game, for example the VR LaboSafe Game, intrinsic motivation and engagement of the trainees are increased significantly. They believe that they are more active, can keep their attention better and enjoy the realisation of relevant situations in a virtual environment. This leads to a significantly lower controlled motivation and makes people more oriented to higher-ranked motivation profiles.

However, this study also presents limitations of using VR as safety training tool. The digital novelty of VR technology makes it hard for people to get used to such training method. Complicated usability and uncomfortable ergonomics can lead to lower-ranked motivational profiles. Especially older employees (above 50 years old) in this study have more issues with this technology than younger employees (below 30 years old). Moreover, some learning content are better taught in a classroom than with VR. So, people are still unfamiliar with VR serious games being used in safety training which explains that some people believe that safety training with VR serious games is not more important for them than with conventional methods. In order to overcome these shortcomings, it is suggested to combine conventional methods with VR as complementary tool and provide more frequent and smaller sessions, gradually introducing VR technology to beginners.

One of the limitations of this study is that results of the motivational profile analysis could be affected due to the relatively small sample population. However, a larger sample size could not be achieved due to the extensive testing duration and due to the limited availability of volunteering employees at the chemical company. Another limitation of this study is that the elevated engagement and motivation could be subdued by the volunteer bias. The voluntary participation of the employees could mean they already had an increased interest in the VR technology, thus explaining the increased intrinsic motivation. Though, obliging employees to participate would raise ethical issues and would also affect the results.

At the end, employees are more intrinsically motivated and are more willing to reengage with VR safety training than with conventional training methods only. This leads to a better attitude towards safety training and could eventually bring better safety outcomes. Although, participants in this study are highly educated personnel working in chemical laboratories, similar motivational results should be expected for other kinds of safety training, taking into account that the VR training is specifically tailored to their real job.

CHAPTER 6

General Conclusions and Future Perspectives

In this doctoral dissertation a complete process is presented from ideation of a VR safety training to a fully developed and evaluated VR training application. An example is shown of a VR serious game for a chemical lab safety training, but the process of creating and implementing VR serious games can be generalised for other safety training programmes. Furthermore, an evaluation study with this VR serious game reveals the impact of using such training tool on the motivation and engagement of employees in a chemical company environment compared to a more conventional training method. This chapter is intended to give a general overview of results achieved in each chapter and concrete suggestions are provided for future research and exploitation.

6.1 General conclusions

The main scope of this doctoral research project is to investigate the potential value of VR serious games as a novel training method for in-company and academic chemical laboratory safety training. Therefore, this research study applies a design-based research methodology to design and develop a VR serious game with design guidelines that are applicable for other VR safety training games. This game was also evaluated as a solution to solve a real-world problem of improving motivation and engagement for safety training.

In Chapter 2, we analysed the state of the art of virtual chemical laboratories. A systematic literature review was performed to summarise the research, technology and instructional design of virtual chemical laboratories. Papers on virtual labs for educating laboratory practices were also included because of the lack of literature on virtual labs for safety training. Results of this review revealed that virtual labs can provide better results in learning outcomes of all domains (i.e. cognitive, affective and skill-based) than passive teaching methods (e.g. lecture, text and video) and show equal or greater effectiveness compared to hands-on laboratory. Better results are shown when virtual labs and traditional methods are combined. Furthermore, most of the included studies use 3D Desktop technology, while immersive VR technology is trending in the last few years. This review also identified that the majority of the studies have not mentioned any instructional design elements. So from this chapter it can be concluded that there is a need to investigate virtual laboratories using immersive VR technology as a tool to motivate and engage learners to follow lab safety training.

The combination of VR technology and serious game should be ideal to encourage motivation and engagement. However, such VR training game needs to be optimally designed and developed. The design and development of such VR serious game is not easy, as it is highly multidisciplinary and not much information can be found in literature about the entire process. Therefore, the need to define formal design and development instructions is fulfilled in Chapter 3. This process is summarised in Figure 6.1 with multiple perspectives and tips on how to design and develop a VR serious game for safety training. These design principles are based on well-known literature studies on how to manage the cognitive load of the learner with instructional design and how to stimulate intrinsic motivation through serious game design. The RAMP guidelines informs which safety skills are important to be taught in a VR safety training. In addition, VR-specific guidelines are discussed that minimise severe symptoms of simulator sickness by improving immersion and interactivity. In Chapter 3, a selection is presented of inexpensive and easily accessible software development tools in order to develop VR serious games. The most important tools are the game engine Unity3D and 3D modelling tool Blender. Particular for novices, online resources are suggested in the chapter that could help with the use of these software tools. Furthermore, development techniques are described in order to optimise the performance of heavily rendered VR environments. The most common practices are polygon count reduction and occlusion culling. Eventually, these design instructions and development suggestions are applied to establish VR LaboSafe Game and can be seen as a illustration of the good practices that are put forward.

In Chapter 4, the VR LaboSafe Game is implemented and tested on an academic population (i.e. from KU Leuven) and on an industrial population (i.e. from Arkema). These tests measured the system usability and simulator sickness when playing the VR game. It is important to ensure that the VR LaboSafe Game is easy-to-use and induces minimal simulator sickness symptoms, because bad usability and severe simulator sickness have a negative influence on learning and motivation. A first test was performed with the VR LaboSafe Game version 1.0 that contained textual instructions in tutorial levels and no animated characters. The participants in this test played the game for 40 minutes without a break. Results from this test revealed that players had difficulties with controlling VR by following textual instructions and that they experience some simulator sickness symptoms. This suggested that adjustments should be implemented to improve system usability and to lower simulator sickness. Therefore, in the updated VR LaboSafe Game version 2.0, textual information was reduced, spoken instructions were included and guiding character animations were introduced. Then, a second test was performed with VR LaboSafe Game 2.0 where the participants were allowed to take breaks from VR between game levels. Results from the second test indeed revealed that game version 2.0 had a superior system usability and lower simulator sickness than version 1.0. A possible explanation is that the cognitive load of the users could be overwhelmed by the high amount of visual information in the game version where textual instructions were displayed. This cognitive overload then has a negative effect on the system usability. Therefore, the solution is to replace the textual information with spoken instructions in order to improve the system usability. In addition, allowing frequent breaks from VR helps the users to rest their eyes, thus, lowering the simulator sickness symptoms. The VR LaboSafe Game version 2.0 is then suitable to be used for further evaluation studies.

In Chapter 5, an evaluation study is described to answer the research question on what the impact is on motivation and engagement when lab safety training is given with a VR serious game and with a conventional training method. In this case, motivation and engagement to follow safety training with a video lecture and with VR LaboSafe Game were measured with employees of the chemical company Arkema. One of the results of this study reveals that employees are mostly autonomously motivated for safety training, which means that they find it important for the safety of themselves and others. With conventional training methods, however, they are less intrinsically motivated. This means that they are not completely motivated by their own and that they find these training courses not entertaining. The reason for that is that they are not active and find it repetitive, thus losing their attention easily. The intrinsic motivation and engagement is increased when VR LaboSafe Game was provided as training tool. Employees are more active during the VR training and enjoy that situations, which are relevant to their job, are realised in the virtual environment. However, one of the limitations is that people do not have the habit of using the novel technology of VR as a learning tool. The VR headset devices can be difficult to use for certain people. More particularly employees of older age (above 50 years old) have more difficulties with handling the VR devices than younger employees (below 30 years old). Suggestions are given on how to overcome the shortcomings of implementing VR serious games for safety training (See Figure 6.1). For example, by offering more frequent and smaller VR sessions while introducing the novel technology gradually, will help trainees familiarise themselves with such technology. Moreover, as discussed in Chapter 2, providing VR serious games as a complementary tool to conventional training methods will make safety training more engaging and effective than those methods separately. In any case, with VR serious games, employees are more willing to follow safety trainings more often than conventional methods. This will lead to a better attitude towards safety training and possibly better safety outcomes. An important remark of this study is that the positive outcomes could be elevated because participants were voluntarily enrolled. There is a possibility that they already have an increased interest in the technology.

As a result of this project, dual outcomes are accomplished. On one hand, a set of guidelines are established based on well-known literature studies and user experiences in order to design, develop and implement VR serious games for safety training (See Figure 6.1). The set of guidelines can contribute as a guide that can be implemented by other researchers, designers and developers who are working in a similar topic. Although, these guidelines were used in the context of chemical lab safety for highly educated personnel, the same design guidelines are applicable for any kind of safety related training (e.g. operator, maintenance, warehouse safety). These guidelines are not limited and can be adjusted depending on the needs of the project. On the other hand, a fully operational VR prototype was created as a usable artefact for chemical laboratory safety training. This real-world application has proven to increase intrinsic motivation and engagement of trainees for safety training. However, additional considerations should then be taken to let people familiarise with the VR technology before fully deploying it in training programmes.

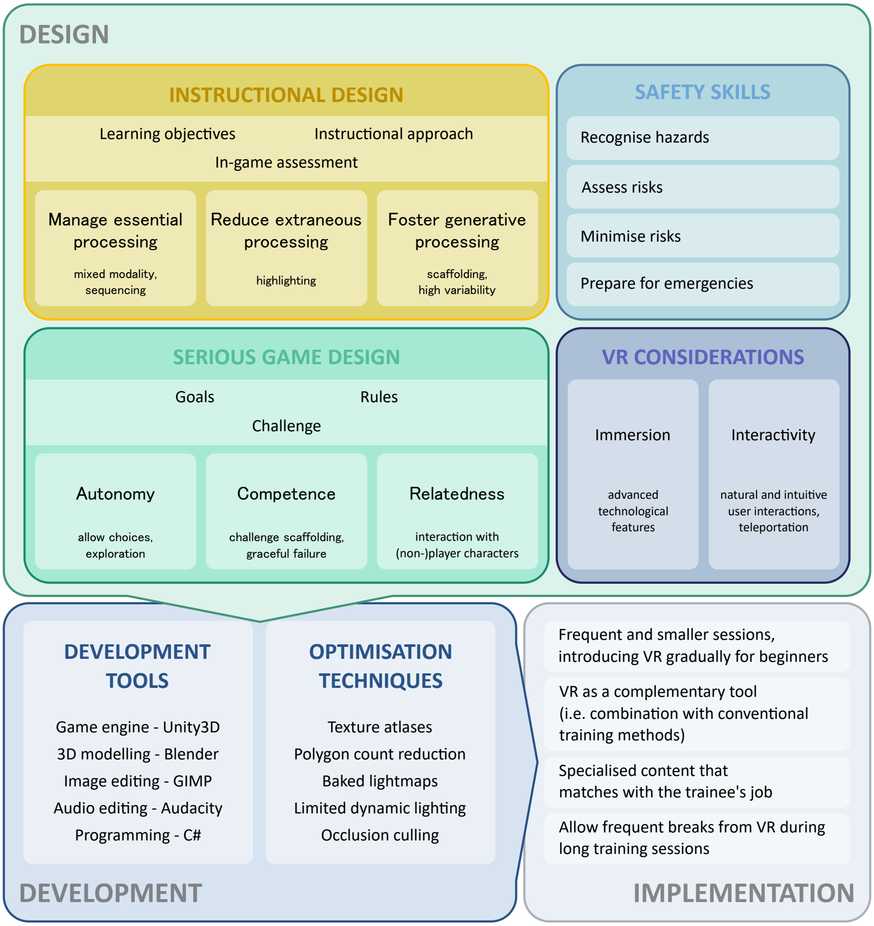


Figure 6.1: Guidelines for the design, development and implementation of VR serious games of safety training programmes.

6.2 Future perspectives

Future perspectives are outlined in terms of research as well as possible exploitation.

This research project performed evaluations on motivation and engagement which belong to learning outcomes in the affective domain. It is concluded that employees are motivated and engaged to play this VR serious game, but it is possible that learning with this tool is ineffective. The design of the VR LaboSafe Game has incorporated elements of the cognitive instructional design principles to optimise learning processes and based on accomplishing safety-related learning objectives. However, no investigations have been made in this research on the effectiveness concerning knowledge and skills acquisition of a VR safety training. The time and effort required to perform such studies are beyond the scope of this research project. Therefore, the next step of this research is to investigate other domains (i.e. knowledge and skills domain) of the effectiveness of using VR serious games for safety training so that further design choices could be made to improve the game. In particular, future research should focus on a thorough analysis on the effectiveness of acquiring laboratory safety skills by means of training with VR serious games. Evaluations can be performed testing the safety awareness and safety behaviour of trainees by providing quizzes or using real-time measurements. Moreover, research can focus on acquisition of procedural skills with VR simulated lab experiments. Then, a comparison can be made with real laboratory activities. Finally, long-term studies with VR serious games as complementary tool for conventional safety training will reveal their potential value of transferring acquired skills to the workplace leading to substantial return of investments for industries. However, in order to perform these studies on learning effectiveness, a detailed and rigorous research plan is required. In this case, not only the reliability and validity of the learning measurement tools should be optimised, but also the reliability and validity of the in-game assessment elements. This is still needed in order to obtain consistent and accurate outcomes.

In recent years, immersive display technologies, such as AR and VR, have experienced a rapid expansion in popularity. This growth emerged mainly because VR/AR devices are now more affordable and more accessible, while improving its performance (Report Linker, 2022). Particularly, the COVID-19 pandemic crisis has substantially increased the interest in such technology; not only for entertainment but also for other use cases, such as fitness, business collaborations, – but most importantly – distance learning and virtual in-person training (Petrock, 2021). The expanding potential value of these technologies has already been noticed by Big Tech companies, such as Facebook (now Meta)

(Kastrenakes & Heath, 2021), Apple (Owen, 2022), Samsung (Benyossef, 2022) and Microsoft (Novet, 2021), which have recently put large investments in VR/AR technologies. Even large chemical enterprises are now considering these technologies as integral part of Industry 4.0 developments; companies such as BASF (Aveva, 2021), Merck (Merck Group, 2021) and Arkema (Arkema, 2019), who are also collaborating partners of the ETN-CHARMING project.

It is evident that industries across all sectors are in urgent need of more advanced and tailor-made VR/AR products. The chemical sector is not an exception since this sector is in high demand of tools to educate and train employees in a virtual and safe environment. Therefore, the next logical step is to take advantage of the knowledge, skills and expertise acquired in this research project in order to establish a VR/AR start-up to support this high demand.

Appendix A

Appendix to chapter 2

Table A.1: A list of categories and their description used for coding the reviewed publications

Variables	Category	Description
Basic publication information		Author, title, publication year
Research purpose	Comparative study	Studies that investigated two or more intervention groups either comparing the media (media comparison) or the design of the virtual laboratory (value-added research) (Mayer, 2014a)
	Evaluative study	Studies that only considered the virtual lab intervention group to evaluate performance assessment, user study or correlation study
	Technical study	Studies that have not performed measurements but describes the design and development of the virtual chemical laboratory
Sample size		Number of participants that were involved in the study
Sample population	Elementary school	Children until 11 years old in elementary school
	Middle/High school	Students between 11 to 18 years old in middle or high school
	University	Students between 18-24 years old in university
	Teachers	Adults older than 24 years old working as teachers

Variables	Category	Description
Comparison	Virtual lab vs passive media vs hands-on lab (or a combination)	The comparison between a virtual chemical laboratory with passive media (e.g., classroom lectures, video, text manual, demonstrations) or with traditional hands-on laboratory, including a combination of these media (e.g., virtual lab + hands-on lab or virtual lab + passive media).
Evaluation method	Test	A quiz testing cognitive outcomes of the student right before (pre-test) or/and after (post-test) the intervention.
	Experiment	A chemical experiment performed in a traditional hands-on laboratory as evaluation of the student's laboratory skills
	Real-time assessment	Embedded performance evaluation within the digital application and recorded in data log files.
	School grade	The school grade of the student after an academic period (usually a trimester or semester)
	Questionnaire	A survey with questions to be answered by the participant
	Interview	A structured conversation with the participant containing questions to be answered
	Observation	Findings by direct observation of the participant during the intervention
Learning outcome	Cognitive	Learning outcomes of the cognitive dimension (including declarative, procedural and conditional knowledge)
	Affective	Learning outcomes of the affective dimension (including self-efficacy, attitude, usability)
	Skill-based	Learning outcomes of the skill-based dimension (including laboratory handling skills)
Technology type	2D Desktop	Two dimensional representation of the lab environment or equipment using a desktop monitor display
	3D Desktop	Three dimensional representation of the lab environment or equipment using a desktop monitor display
	Immersive VR	Immersive virtual reality device that allows a high level of immersion (including head-mounted displays)
	NUI (2D, 3D or imVR)	Natural user interfaces that uses ergonomic movements or gestures as input to control the virtual lab (including spatial tracking, hand gestures, body gestures)

Variables	Category	Description
Instructional design	Instructional approach	Learning theories applied in the virtual lab
	Instructional support	Instructional support elements that serve as an aid for the user's cognitive processing in the virtual lab

Table A.2: Coding scheme of the 76 reviewed publications in the systematic literature review.

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
1	**(Achuthan & Murali, 2015)**	2D	C (media)	141	University	Cog (dk)	test	virtual vs virtual + hands-on vs hands-on	virtual = virtual + hands-on hands-on (dk)	Not specified	Not specified
2	**(Agbonifo et al., 2020)**	3D	E (user study)	50	Middle/High, University	Aff (us)	quest	n/a		Not specified	Not specified
3	**(Aldosari & Marocco, 2015)**	NUI (3D)	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
4	**(Aldosari & Marocco, 2016)**	NUI (3D)	E (user study)	90	University	Aff (us)	quest	n/a		Learning-by-doing	Not specified
5	**(Ali et al., 2014)**	NUI (3D)	C (media)	14	Middle/High	Cog (dk), Aff (us, se), Skill	test, lab pr, quest	virtual vs passive	Virtual passive (dk, sb)	Not specified	modality
6	**(Al-Khalifa, 2017)**	NUI (2D)	E (user study)	16	Elementary	Aff (us)	quest, obs	n/a		Not specified	Not specified
7	**(Almazaydeh et al., 2016)**	2D	E (user study)	25	Middle/High	Aff (us)	quest	n/a		Not specified	Not specified
8	**(Alqadri, 2018)**	2D	E (performance)	30	Middle/High	Cog (dk, pk)	test	n/a		Direct instruction	Not specified
9	**(Annetta et al., 2014)**	3D	E (performance)	31	Teachers	Cog (dk), Other	test, quest, interv	n/a		Problem-based learning	Pedagogical agent, feedback, narrative
10	**(Astuti et al., 2019)**	imVR	C (media)	96	Middle/High	Aff (att)	quest, obs	virtual vs virtual + hands-on vs hands-on	virtual + passive = virtual + hands-on hands-on (att)	Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
11	**(Bakar et al., 2013)**	3D	C (media)	61	Middle/High	Cog (ck)	test, interv, obs	virtual vs passive	Virtual passive (ck)	Cognitivism-constructivism-contextual approach	Not specified
12	**(Bell, 2004)**	3D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
13	**(Borek et al., 2009)**	2D	C (value-added)	87	University	Cog (dk, ck), Aff (us)	test, quest	Inquiry vs tutored vs direct instruction	Tutored condition = inquiry = direct instruction (dk) Tutored condition > inquiry or direct-instruction (ck)	Inquiry learning	feedback, scaffolding guidance
14	**(Chee & Tan, 2012)**	3D	C (media)	39	Middle/High	Cog (dk pk), Aff (att)	test, quest, interv, obs	virtual vs passive	virtual passive (dk, pk)	Inquiry-based, learning-by-doing, embodied learning	narrative
15	**(Chee & Tan, 2012)**	2D	E (user study)	428	University	Aff (us)	quest, interv	n/a		Not specified	Modality, personalisation
16	**(Cuadros et al., 2015)**	2D	E (performance)	60	Middle/High	Cog (pk)	r-t assess	n/a		Not specified	n/a
17	**(Dalgarno & Lee, 2009)**	3D	C (media)	133	University	Cog (dk), Aff (att, us, se)	test, quest, interv	virtual vs hands-on	virtual = hands-on (dk, att, se)	Not specified	
18	**(Davenport et al., 2018)**	2D	C (media)	1334	Middle/High	Cog (dk, pk)	test, interv, r-t assess	virtual vs virtual + passive	virtual + passive virtual (dk)	Not specified	Scaffolding guidance
19	**(Desai et al., 2017)**	NUI (3D)	E (performance, user study)	31	Middle/High	Aff (us), Skill	quesst, r-t assess	n/a		Not specified	Scaffolding feedback

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
20	**(N. D. Dholakiya et al., 2019)**	3D	E (user study)	45	University	Aff (us)	quest	n/a		Kolb's experiential learning	Not specified
21	**(N. S. Dholakiya et al., 2019)**	3D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
22	**(Donnelly et al., 2013)**	2D	E (user study)	4	Teachers	Aff (att)	quest, interv, obs	n/a		Inquiry-based learning	Scaffolding guidance, feedback
23	**(Duan et al., 2020)**	imVR	E (user study)	45	University	Aff (us)	quest, interv	n/a		Not specified	n/a
24	**(Dunnagan et al., 2020)**	imVR	C (media)	75	University	Cog (dk)	test	virtual vs hands-on	virtual = hands-on (dk)	Not specified	Pedagogical agent
25	**(Enneking et al., 2019)**	3D	C (media)	1141	University	Cog (dk, pk, ck), Aff (att), Skill	test, quest, lab pr	virtual + hands-on vs hand-on	virtual + hands-on = hands-on (dk,pk,ck,sb) virtual + hands-on < hands-on (att)	Not specified	Not specified
26	**(Gal et al., 2015)**	2D	E (performance)	306	University	Cog (pk, ck)	r-t assess	n/a		Learning-by-doing, inquiry learning	feedback, scaffolding
27	**(Georgiou et al., 2007)**	3D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
28	**(Gervasi et al., 2004)**	3D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
29	**(Han et al., 2017)**	NUI (imVR)	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
30	**(Hawkins & Phelps, 2013)**	2D	C (media)	169	University	Cog (dk), Skill	test, lab pr	virtual vs hands-on	virtual = hands-on (dk, sb)	Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
31	**(Hensen & Barbera, 2019)**	3D	C (media)	396	University	Aff (att, us)	quest	virtual vs hands-on	virtual hands-on (att)	Not specified	Not specified
32	**(Hensen et al., 2020)**	3D	C (media)	717	University	Cog (dk, ck), Aff (att), Skill	test, quest	virtual vs hands-on	virtual = hands-on (dk, ck, att)	Not specified	Not specified
33	**(Herga & Dinevski, 2012)**	2D	C (media)	38	Elementary	Cog (dk, ck)	test	virtual vs passive	virtual passive (dk,sk)	Not specified	Not specified
34	**(Herga et al., 2015)**	2D	C (media)	225	Elementary	Cog (dk, ck)	test	virtual vs passive	Virtual = passive (dk) Virtual > passive (ck)	Not specified	Not specified
35	**(Herga et al., 2016)**	2D	C (media)	109	Elementary	Cog (dk, ck)	test	virtual vs passive	Virtual b passive (dk, ck)	Not specified	Not specified
36	**(Ikhsan et al., 2020)**	imVR	C (media)	96	Middle/High	Cog (dk)	test	virtual vs virtual + hands-on vs hands-on	virtual + passive = virtual + hands-on hands-on (ck)	Not specified	Not specified
37	**(Ikram et al., 2015)**	NUI (3D)	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
38	**(Irby et al., 2018)**	2D	C (media)	67	University	Cog (dk, ck)	test, obs	Virtual + hands-on vs hand-on	virtual + hands-on = hands-on (dk,ck)	Not specified	Not specified
39	**(Jagodzinski & Wolski, 2015)**	NUI (2D)	C (media, value-added)	200	Middle/High	Cog (dk, ck), Aff (us, se)	test, quest	Media: virtual + passive vs passive Value-added: voice vs text vs video instructions	Media: virtual + passive > passive (dk, ck) Value added: voice > text = video (dk,ck)	Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
40	**(Jagodzinski & Wolski, 2014)**	NUI (2D)	C (media)	150	Middle/High	Cog (dk, ck), Aff (us, se)	test, quest, obs	virtual vs passive	Virtual > passive (sk) virtual = passive (dk)	Embodied cognition	Not specified
41	**(Jordá, 2013)**	3D	E (user study)	15	University	Aff (us)	quest	n/a		Not specified	Not specified
42	**(Y. Kim et al., 2016)**	NUI (3D)	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
43	**(Y. Kim et al., 2017)**	NUI (3D)	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
44	**(H. Kim et al., 2019)**	imVR	T	n/a	n/a	n/a	n/a	n/a		Not specified	Scaffolding guidance, feedback
45	**(Kolil et al., 2020)**	2D	C (media)	1225	University	Aff (se)	quest	virtual + hands-on vs hands-on	virtual + hands-on hands-on (se)	Not specified	Not specified
46	**(Lau, 2017)**	3D	T	n/a	n/a	n/a	n/a	n/a		Situated learning	Not specified
47	**(Makransky, Borre-Gude, & Mayer, 2019)**	imVR	C (media)	105	University	Cog (dk, ck), Aff (se), Skill	test, lab pr, quest, interv	virtual vs passive	virtual = passive (dk) virtual > passive (ck)	Control value theory, embodied cognition	Pedagogical agent, modality, feedback
48	**(Makransky, Wismer, & Mayer, 2019)**	imVR	C (value-added)	66	Middle/High	Cog (dk, ck), Aff (social)	test, quest	Human female vs robot drone	Human female > robot drone (for girls) (dk,ck) Robot drone > human female (for boys) (dk, ck)	Not specified	Pedagogical agent, modality, feedback
49	**(Martinez-Jimenez et al., 2003)**	2D	C (media)	274	University	Cog(dk, pk, ck), Skill	test, lab pr	virtual + hands-on vs hands-on	Virtual + hands-on hands-on (dk,pk,ck,sb)	Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
50	**(Moozeh et al., 2020)**	2D	E (user study)	46	University	Aff (us)	quest	n/a		Not specified	feedback
51	**(Morozov et al., 2004)**	3D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Pedagogical agent
52	**(Penn & Ramnarain, 2019)**	2D	E (user study)	50	University	Aff (att, us, se)	quest, interv	n/a		Not specified	Not specified
53	**(Pyatt & Sims, 2012)**	2D	C (media)	184	Middle/High	Cog (dk, pk), Aff (us,att), Skill	lab pr, quest	virtual vs hands-on	trial 1: virtual = hands-on (dk,pk) trial 2: virtual > hands-on (dk,pk)	Inquiry-based, guided-discovery, learning-by-doing	Not specified
54	**(Qvist et al., 2015)**	3D	E (user study)	29	University	Aff (us)	quest, interv, obs			learning-by-doing	Not specified
55	**(Ramos et al., 2016)**	2D	E (user study)	120	University	Aff (us)	quest	n/a			Not specified
56	**(Ratamun & Osman, 2018)**	2D	C (media)	147	Middle/High	Cog (dk, ck)	test	virtual vs hands-on	virtual hands-on (dk,ck)	inquiry-based learning	Not specified
57	**(Rowe et al., 2018)**	2D	C (media)	160	University	Cog (dk), Aff (us), Skill	quest, grade	1: virtual vs hands-on 2: virtual + passive vs passive	1: virtual = hands-on (dk, sb) 2: virtual + passive > passive (dk)	inquiry-based learning	Not specified
58	**(Sampaio et al., 2014)**	3D	E (performance, user study)	6	Middle/High	Cog (pk), Aff (us)	quest, obs, interv, r-t assess	n/a		Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
59	**(Scherer & Tiemann, 2012)**	2D	E (other)	162	Middle/High	Cog (ck)	quesst, r-t assess	n/a		Problem-solving based design	Not specified
60	**(Solikhin et al., 2019)**	2D	C (media)	87	Middle/High	Cog(dk), Aff (us)	test, quest	virtual vs virtual + hands-on vs hands-on	virtual + hands-on virtual hands-on (dk)	Not specified	Not specified
61	**(Su & Cheng, 2019)**	3D	C (media)	72	Middle/High	Cog (dk), Aff (se)	test, quest	virtual vs passive	virtual passive (dk)	Kolb's experiential learning, cognitive load theory	Feedback, scaffolding guidance
62	**(Tarng et al., 2017)**	3D	C (media)	80	University	Cog (dk), Aff (us)	test, quest	virtual vs hands-on	Virtual hands-on (dk)	Situated learning	Not specified
63	**(Tatli & Ayas, 2012)**	3D	C (media)	90	Middle/High	Other	quest, interv, obs	virtual vs passive + hands-on vs passive	virtual passive + hands-on passive	predict-observe-explain (POE)	Not specified
64	**(Tatli & Ayas, 2013)**	3D	C (media)	90	Middle/High	Cog (dk)	test, obs, interv	virtual vs hands-on	Trial 1: virtual = hands-on (dk) Trial 2: virtual > hands-on (dk)	Constructivist approaches	Not specified
65	**(Ullah et al., 2016)**	NUI (3D)	C (media, value-added)	57	Middle/High	Cog (dk, pk), Aff (us, se), Skill	test, quest, r-t assess, lab pr	Media: virtual vs hands-on Value-added: procedural vs non-procedural	Media: virtual > hands-on (dk, sb) Value-added: procedural > non-procedural (sb)	Inquiry-based, discovery based learning	Modality, scaffolding guidance

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
66	**(Wang & Dong, 2018)**	imVR	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified
67	**(Winkelmann et al., 2014)**	3D	C (media)	12	Middle/High	Cog (dk, ck), Aff (att, us)	test, quest	virtual vs hands-on	virtual = hands-on (dk,ck, att)	Not specified	Not specified
68	**(Winkelmann et al., 2017)**	3D	C (media)	122	University	Cog (dk), Aff (us), Skill	test, quest, interv, obs, lab pr	virtual vs hands-on	virtual = hands-on (dk, sb)	Not specified	Not specified
69	**(Winkelmann et al., 2020)**	3D	C (media)	279	University	Cog (dk), Aff (att, us), Skill	test, quest, lab pr	virtual + hands-on vs hands-on	virtual + hands-on = hands-on (dk, att, sb)	Not specified	Not specified
70	**(Wolski & Jagodzinski, 2019)**	NUI (2D)	C (media)	150	Middle/High	Cog (dk, ck)	test	virtual vs passive	virtual = passive (dk) virtual > passive (ck)	Not specified	Not specified
71	**(Woodfield et al., 2004)**	3D	E (user study)	616	University	Aff (us, se)	quest, interv, obs	n/a		Not specified	Not specified
72	**(Woodfield et al., 2005)**	3D	E (performance)	963	University	Cog (dk, ck), Aff (us)	quest, interv, obs, grade	n/a		Not specified	Not specified
73	**(Wu et al., 2019)**	NUI (imVR)	E (performance, user study)	28	University	Aff (us), Skill	quest, r-t assess	n/a		Not specified	Not specified
74	**(Yaron et al., 2010)**	2D	T	n/a	n/a	n/a	n/a	n/a		Not specified	Not specified

No.	Author	Tech	Method	Size	Population	Learning outcome	Evaluation method	Comparison	Media comparison	Learning theory	Instructional support
75	**(Zayas-Pérez & Cox, 2009)**	3D	C (value-added)	48	University	Cog (dk)	test	Co-located vs not co-located	Co-located not co-located (dk)	Learning-by-doing	spatial contiguity, feedback, modality
76	**(Zhong & Liu, 2014)**	3D	E (user study)	14	Teachers	Aff (us)	quest, interv	n/a		Not specified	Not specified

Abbreviations: 2D = 2D Desktop; 3D = 3D Desktop; imVR = immersive VR; NUI = Natural User Interface; C = comparative; E = evaluative; T = technical, Cog = cognitive domain, Aff = affective domain, Skill = skill-based domain, dk = declarative knowledge; pk = procedural knowledge; ck = conditional knowledge; att = attitude; se = self-efficacy; us = usability; quest = questionnaire; interv = interview; r-t assess = real-time assessment; obs = observation; lab pr = lab practical; n/a = not available.

Appendix B

Appendix to chapter 4

Table B.1: Adapted items from the SUS questionnaire (Brooke, 1995). Even numbered items (^{R}) are reverse-coded

Item number	SUS item
S1	I think that I would like to use the VR LaboSafe Game frequently
S2^{R}	I found the VR LaboSafe Game unnecessarily complex
S3	I thought the VR LaboSafe Game was easy to use
S4^{R}	I think that I would need the support of a technical person to be able to use the VR LaboSafe Game
S5	I found that the various functions in the VR LaboSafe Game were well integrated
S6^{R}	I thought that there was too much inconsistency in the VR LaboSafe Game
S7	I would imagine that most people would learn to use the VR LaboSafe Game very quickly
S8^{R}	I found the VR LaboSafe Game very awkward to use
S9	I felt very confident using the VR LaboSafe Game
S10^{R}	I needed to learn a lot of things before I could get going with the VR LaboSafe Game

Appendix C

Appendix to chapter 5

Table C.1: Questionnaire items on the attributes of engagement.

Subscale	**Items**	Source
Absorption in the task	E1. This safety training (with VR serious games) was not boring for me	(Magyaródi et al., 2013)
	E2. Following this safety training (with VR serious games) totally engrossed my attention	
	E3. I forgot about the progress of time	
	E4. Time passed faster than I thought	
Control active learning	E5. This safety training (with VR serious games) allows me to be more responsive and active in the learning process.	(Lee et al., 2010)
	E6. This safety training (with VR serious games) allows me to have more control over my own training of safety awareness and safety behaviour.	
	E7. This safety training (with VR serious games) promotes self-paced learning.	
	E8. This safety training (with VR serious games) helps to get myself engaged in safety trainings.	
Behavioural Intention	E9. I intend to follow this safety training (with VR serious games) more, assuming I had access to it for a relevant subject	(Makransky & Lilleholt, 2018)
	E10. I would follow this safety training (with VR serious games) frequently in the future	
	E11. I would like to participate in other safety trainings that uses this teaching method (with VR serious games)	
	E12. I would train my safety skills more if I had access to this safety training (with VR serious games) in my workplace	

Table C.2: Questionnaire items on safety training motivation (adapted from (Scott, 2016)).

Subscale	Items
Intrinsic motivation	M1. Because I take pleasure in following safety trainings (with VR serious games)
	M2. Because safety trainings (with VR serious games) are fun
	M3. Because safety trainings (with VR serious games) interest me
	M4. Because I enjoy following safety trainings (with VR serious games)
Identified regulation	M5. Because I believe it is important to follow safety trainings (with VR serious games)
	M6. Because safety trainings (with VR serious games) are personally important to me
	M7. Because I want to learn new things
	M8. Because I believe safety trainings (with VR serious games) improve my safety skills
Introjected regulation	M9. Because I would feel ashamed of myself if I didn't follow safety trainings (with VR serious games)
	M10. Because otherwise I will feel guilty
	M11. Because I want others (e.g., supervisor, colleagues, family, client) to think I take safety trainings (with VR serious games) seriously
	M12. Because I want others (e.g., supervisor, colleagues, family, client) to think I have good safety skills
External regulation	M13. Because others (e.g., supervisor, colleagues, family, client) oblige me to follow safety trainings (with VR serious games)
	M14. In order to get approval from others (e.g., supervisor, colleagues, family, client)
	M15. Because I am supposed to follow safety trainings (with VR serious games)
	M16. Because I risk losing my job if I don't
	M17. In order to acquire an attendance certificate for following safety trainings (with VR serious games)
Amotivation	M18. I am not motivated, because it doesn't make a difference whether I follow safety trainings (with VR serious games) or not.
	M19. I am not motivated, because safety trainings (with VR serious games) are not a priority to me
	M20. I am not motivated, because safety trainings (with VR serious games) are not a priority in my workplace
	M21. I am not motivated, because following safety trainings (with VR serious games) are not worth the effort

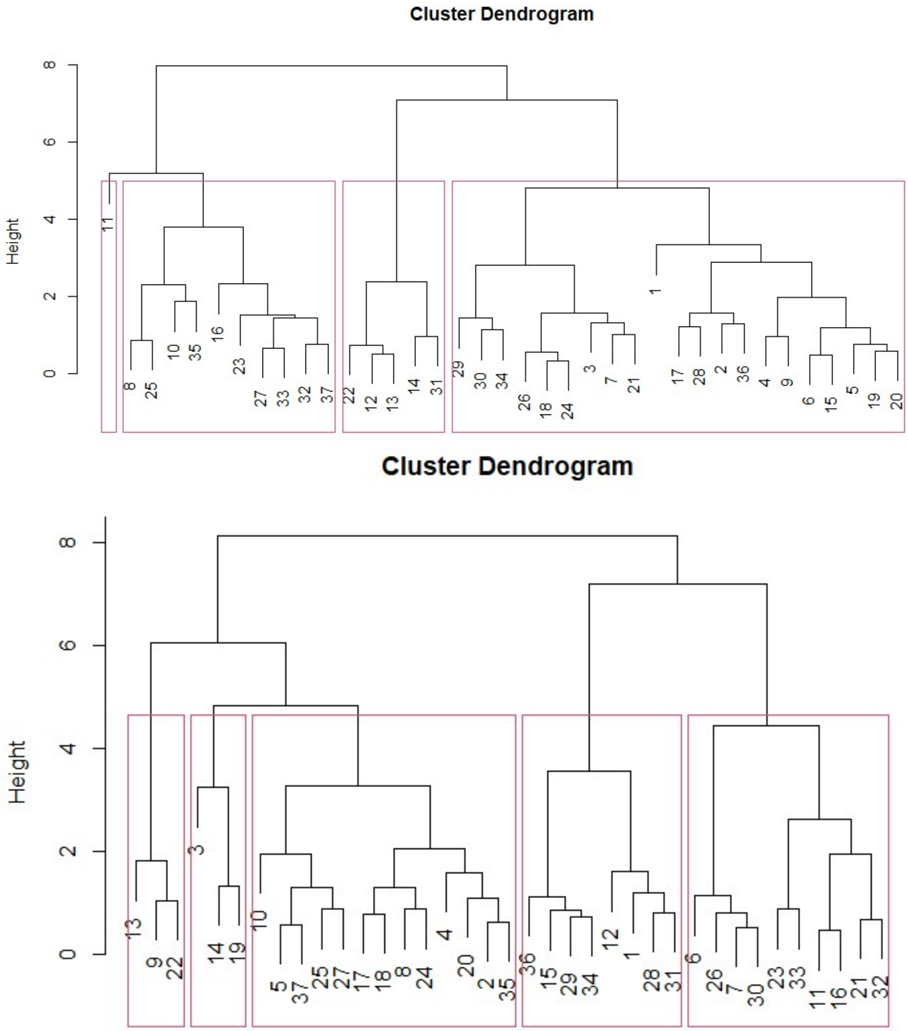


Figure C.1: Dendrogram representations of the hierarchical clustering method of (top) pre-test and (bottom) post-test motivation data.

Bibliography

Achuthan, K., & Murali, S. S. (2015). A Comparative Study of Educational Laboratories from Cost & Learning Effectiveness Perspective. In Silhavy, R and Senkerik, R and Oplatkova, ZK and Prokopova, Z and Silhavy, P (Ed.), *Software engineering in intelligent systems (csoc2015), volume 3* (Vol. 349, p. 143-153). (4th Computer Science On-line Conference (CSOC), CZECH REPUBLIC, APR 27-30, 2015) doi: 10.1007/978-3-319-18473-9_15

Agbonifo, O. C., Sarumi, O. A., & Akinola, Y. M. (2020). A chemistry laboratory platform enhanced with virtual reality for students' adaptive learning. *Research in Learning Technology*, *28*. doi: 10.25304/rlt.v28.2419

Akçayır, M., Dündar, H., & Akçayır, G. (2016). What makes you a digital native? is it enough to be born after 1980? *Computers in Human Behavior*, *60*, 435–440. doi: 10.1016/j.chb.2016.02.089

Alaimo, P. J., Langenhan, J. M., Tanner, M. J., & Ferrenberg, S. M. (2010). Safety teams: An approach to engage students in laboratory safety. *Journal of Chemical Education*, *87*(8), 856-861. doi: 10.1021/ed100207d

Aldosari, S. S., & Marocco, D. (2015). Using Haptic Technology for Education in Chemistry. In *Proceedings 2015 fifth international conference on e-learning (econf 2015)* (p. 58-64). (5th International Conference on e-Learning (econf), Manama, BAHRAIN, OCT 18-20, 2015) doi: 10.1109/ECONF.2015.25

Aldosari, S. S., & Marocco, D. (2016). Experiments on Virtual Manipulation in Chemistry Education. In Verbert, K and Sharples, M and Klobucar, T (Ed.), *Adaptive and adaptable learning, ec-tel 2016* (Vol. 9891, p. 543-546). (11th European Conference on Technology-Enhanced Learning (EC-TEL), Lyon, FRANCE, SEP 13-16, 2016) doi: 10.1007/978-3-319-45153-4_59

Ali, N., & Ullah, S. (2020). Review to analyze and compare virtual chemistry laboratories for their use in education. *Journal of Chemical Education*, *97*(10), 3563-3574. doi: 10.1021/acs.jchemed.0c00185

Ali, N., Ullah, S., Rabbi, I., & Alam, A. (2014). The Effect of Multimodal Virtual Chemistry Laboratory on Students' Learning Improvement. In

DePaolis, LT and Mongelli, A (Ed.), *Augmented and virtual reality, avr 2014* (Vol. 8853, p. 65-76). (1st International Conference on Augmented and Virtual Reality (AVR), Lecce, ITALY, SEP 17-20, 2014) doi: 10.1007/978-3-319-13969-2_5

Alkhaldi, T., Pranata, I., & Athauda, R. I. (2016). A review of contemporary virtual and remote laboratory implementations: observations and findings. *Journal of Computers in Education*, *3*(3), 329–351. doi: 10.1007/s40692-016-0068-z

Al-Khalifa, H. S. (2017). CHEMOTION: A gesture based chemistry virtual laboratory with leap motion. *Computers Applications in Engineering Education*, *25*(6), 961-976. doi: 10.1002/cae.21848

Almazaydeh, L., Younes, I., & Elleithy, K. (2016). An Interactive and Self-instructional Virtual Chemistry Laboratory. *International Journal of Emerging Technologies in Learning*, *11*(7), 70-73. doi: 10.3991/ijet.v11i07.5853

Alqadri, Z. (2018). Using virtual laboratory in direct instruction to enhance students' achievement. In Uslu, F (Ed.), *5th international conference on education and social sciences (inctcess 2018)* (p. 792-800). (5th International Conference on Education and Social Sciences (INTCESS), Istanbul, TURKEY, FEB 05-07, 2018)

Amokrane, K., & Lourdeaux, D. (2009, July). Virtual Reality Contribution to training and risk prevention. In C. Press (Ed.), *ICAI 2009: the International Conference on Artificial Intelligence 2009* (p. 466-472). Las Vegas, United States. Retrieved from https://hal.archives-ouvertes.fr/hal-00462120

Anderson, C. G., Dalsen, J., Kumar, V., Berland, M., & Steinkuehler, C. (2018). Failing up: How failure in a game environment promotes learning through discourse. *Thinking Skills and Creativity*, *30*, 135–144. doi: 10.1016/j.tsc.2018.03.002

Annetta, L., Lamb, R., Minogue, J., Folta, E., Holmes, S., Vallett, D., & Cheng, R. (2014). Safe science classrooms: Teacher training through serious educational games. *Information Sciences*, *264*, 61-74. doi: 10.1016/j.ins.2013.10.028

Arkema. (2019). *Innovative - 2019 annual and sustainable performance report.* Retrieved from https://e-brochure.arkema.com/media/2019-annual-sustainable-performance-report/article/62/

Arkema. (2022a). *Acting as a responsible manufacturer: Integrate safety as a must.* Retrieved from https://www.arkema.com/global/en/social-responsibility/responsible-manufacturer/industrial-safety/

Arkema. (2022b). *Arkema, a leader in specialty materials.* Retrieved from https://www.arkema.com/global/en/arkema-group/profile/

Arnab, S., & Clarke, S. (2017). Towards a trans-disciplinary methodology for a game-based intervention development process. *British Journal of Educational Technology*, *48*(2), 279-312. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12377 doi: 10.1111/bjet.12377

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., ... De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*, *46*(2), 391-411. doi: 10.1111/bjet.12113

Ashwin, P., & McVitty, D. (2015). The meanings of student engagement: Implications for policies and practices. In *The european higher education area* (pp. 343–359). Springer International Publishing. doi: 10.1007/978-3-319-20877-0_23

Astuti, T. N., Sugiyarto, K. H., & Ikhsan, J. (2019). Using virtual reality towards students' scientific attitude in chemical bonding. *European Journal of Education Studies*, *6*(2), 224-238. doi: 10.5281/ZENODO.2958411

Aveva. (2021). *How basf prepares new operators for a sustainable chemicals industry.* Retrieved from https://www.aveva.com/en/perspectives/success-stories/basf/

Bakar, N., Badioze Zaman, H., Kamalrudin, M., Jusoff, K., & Khamis, N. (2013). An effective virtual laboratory approach for chemistry. *Australian Journal of Basic and Applied Sciences*(7 (3)).

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *J. Usability Studies*, *4*(3), 114–123.

Bates, T. (2015). *Teaching in a digital age: Guidelines for teaching and learning.* Vancouver: Tony Bates Associates Ltd.

Bauer, C. F. (2008). Attitude toward chemistry: A semantic differential instrument for assessing curriculum impacts. *Journal of Chemical Education*, *85*(10), 1440. doi: 10.1021/ed085p1440

Bell, J., & Fogler, H. (2004). The application of virtual reality to (chemical engineering) education. In Ikel, Y and Gobel, M and Chen, J (Ed.), *Ieee virtual reality 2004* (p. 217-218). (IEEE Virtual Reality 2004 Conference, Chicago, IL, MAR 27-31, 2004) doi: 10.1109/VR.2004.1310077

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of serious games: An overview. *Advances in Human-Computer Interaction*, *2013*, 1–11. doi: 10.1155/2013/136864

Bellou, I., Papachristos, N. M., & Mikropoulos, T. A. (2018). Digital learning technologies in chemistry education: A review. In D. Sampson, D. Ifenthaler, J. M. Spector, & P. Isaías (Eds.), *Digital technologies: Sustainable innovations for improving teaching and learning* (pp. 57–80). Cham: Springer International Publishing. doi: 10.1007/978-3-319-73417-0_4

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology*, *53*, 87–105. doi: 10.1016/j.cedpsych.2018.01.002

Benyossef, R. (2022). Why we invested in liv, the vr streaming and recording platform. *Samsung NEXT*. Retrieved from https://www.samsungnext.com/blog/why-we-invested-in-liv-the-vr-streaming-and-recording-platform

Bhide, S., Riad, R., Rabelo, L., Pastrana, J., Katsarsky, A., & Ford, C. (2015). Development of virtual reality environment for safety training. *IIE Annual Conference.Proceedings*, 2302-2312. Retrieved from https://search.proquest.com/docview/1792029019?accountid=28387

Bhusari, A., Goh, A., Ai, H., Sathanapally, S., Jalal, M., & Mentzer, R. A. (2020). Process safety incidents across 14 industries. *Process Safety Progress*, *40*(1), e12158. doi: 10.1002/prs.12158

Blair, E., & Seo, D.-C. (2007). Safety Training Making the Connection to High Performance. *Professional Safety*, *52*(10).

Blattgerste, J., Behrends, J., & Pfeiffer, T. (2022). *A Web-Based Analysis Toolkit for the System Usability Scale.* Retrieved from https://github.com/jblattgerste/sus-analysis-toolkit

Blender. (2022). *Blender 3d modelling software.* Retrieved 2022-03-15, from https://www.blender.org/

Blender Guru. (2022). *Blender guru's youtube channel.* Retrieved 2022-03-15, from https://www.youtube.com/channel/UCOKHwx1VCdgnxwbjyb9Iu1g

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *International Journal of Educational Technology in Higher Education*, *17*(1). doi: 10.1186/s41239-019-0176-8

Borek, A., McLaren, B. M., Karabinos, M., & Yaron, D. (2009). How Much Assistance Is Helpful to Students in Discovery Learning? In Cress, U and Dimitrova, V and Specht, M (Ed.), *Learning in the synergy of multiple disciplines* (Vol. 5794, p. 391+). (4th European Conference on Technology Enhanced Learning, Nice, FRANCE, SEP 29-OCT 02, 2009) doi: 10.1007/978-3-642-04636-0_38

Bouchard, S., Berthiaume, M., Robillard, G., Forget, H., Daudelin-Peltier, C., Renaud, P., … Fiset, D. (2021). Arguing in favor of revising the simulator sickness questionnaire factor structure when assessing side effects induced by immersions in virtual reality. *Frontiers in Psychiatry*, *12*. doi: 10.3389/fpsyt.2021.739742

Braad, E., Zavcer, G., & Sandovar, A. (2016). Processes and models for serious game design and development. In R. Dörner, S. Göbel, M. Kickmeier-Rust, M. Masuch, & K. Zweig (Eds.), *Entertainment computing and serious games* (pp. 92–118). Springer. doi: 10.1007/978-3-319-46152-6_5

Brackeys. (2022). *Brackey's youtube channel.* Retrieved 2022-03-15, from `https://www.youtube.com/channel/UCYbK_tjZ2OrIZFBvU6CCMiA`

Branch, R. M. (2009). *Instructional design: The ADDIE approach.* Springer US. doi: 10.1007/978-0-387-09506-6

Bretz, S. L. (2019). Evidence for the importance of laboratory courses. *J. Chem. Educ.*, *96*(2), 193–195. doi: 10.1021/acs.jchemed.8b00874

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education*, *87*, 218–237. doi: 10.1016/j.compedu.2015.07.003

Brooke, J. (1995). Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, *189*.

Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American journal of public health*, *96*(16380566), 315–324. doi: 10.2105/AJPH.2004.059840

Buttussi, F., & Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063-1076. doi: 10.1109/TVCG.2017.2653117

C# Documentation. (2022). *C# Documentation.* Retrieved 2022-03-15, from `https://docs.microsoft.com/en-us/dotnet/csharp/`

Caserman, P., Garcia-Agundez, A., Zerban, A. G., & Göbel, S. (2021). Cybersickness in current-generation virtual reality head-mounted displays: systematic review and outlook. *Virtual Reality*, *25*(4), 1153–1170. doi: 10.1007/s10055-021-00513-6

Casey, T., Turner, N., Hu, X., & Bancroft, K. (2021). Making safety training stickier: A richer model of safety training engagement and transfer. *Journal of Safety Research*, *78*, 303–313. doi: 10.1016/j.jsr.2021.06.004

Checa, D., & Bustillo, A. (2019). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, *79*(9-10), 5501–5527. doi: 10.1007/s11042-019-08348-9

Chee, Y. S., & Tan, K. C. D. (2012). Becoming Chemists through Game-based Inquiry Learning: The Case of Legends of Alkhimia. *Electronic Journal of E-learning*, *10*(2), 185-198.

Chen, M., Wu, Y., Wang, K., Guo, H., & Ke, W. (2020). An explosion accident analysis of the laboratory in university. *Process Safety Progress*, e12150. doi: 10.1002/prs.12150

Cherni, H., Métayer, N., & Souliman, N. (2020). Literature review of locomotion techniques in virtual reality. *International Journal of Virtual Reality*, *20*(1), 1–20. doi: 10.20870/ijvr.2020.20.1.3183

Code Monkey. (2022). *Code monkey's youtube channel.* Retrieved 2022-03-15, from `https://www.youtube.com/c/CodeMonkeyUnity`

Crookall, D. (2010). Serious games, debriefing, and simulation/gaming as a discipline. *Simulation & Gaming*, *41*(6), 898–920. doi: 10.1177/1046878110390784

Csikszentmihalyi, M., Abuhamdeh, S., Nakamura, J., et al. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Cuadros, J., Artigas, C., Guitart, F., & Martori, F. (2015). Analyzing a virtual-lab based contextualized activity from action logs. In Ozdamli, F (Ed.), *4th world conference on educational technology researches (wcetr-2014)* (Vol. 182, p. 441-447). (4th World Conference on Educational Technology Researches (WCETR), Univ Barcelona, Barcelona, SPAIN, NOV 28-29, 2014) doi: 10.1016/j.sbspro.2015.04.823

Dakkoune, A., Vernières-Hassimi, L., Leveneur, S., Lefebvre, D., & Estel, L. (2018). Risk analysis of french chemical industry. *Safety Science*, *105*, 77–85. doi: 10.1016/j.ssci.2018.02.003

Dalgarno, B., Bishop, A. G., Adlong, W., & Bedgood, D. R., Jr. (2009). Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. *Computers & Education*, *53*(3), 853-865. doi: 10.1016/j.compedu.2009.05.005

Dalgarno, B., & Lee, M. J. W. (2009). What are the learning affordances of 3-d virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. doi: 10.1111/j.1467-8535.2009.01038.x

Davenport, J. L., Rafferty, A. N., & Yaron, D. J. (2018). Whether and How Authentic Contexts Using a Virtual Chemistry Lab Support Learning. *Journal of Chemical Education*, *95*(8), 1250-1259. doi: 10.1021/acs.jchemed.8b00048

Davis, S., Nesbitt, K., & Nalivaiko, E. (2015). Comparing the onset of cybersickness using the oculus rift and two virtual roller coasters. In *Proceedings of the 11th australasian conference on interactive entertainment (ie 2015)* (Vol. 27, p. 30).

Deci, E. L., & Ryan, R. M. (2004). *Handbook of self-determination research.* Boydell & Brewer Ltd. Retrieved from `https://www.ebook.de/de/product/2764537/edward_l_deci_richard_m_ryan_handbook_of_self_determination_research.html`

Desai, K., Jin, R., Prabhakaran, B., Diehl, P., Hernndez Belmonte, U. H., Ayala Ramirez, V., . . . Gans, M. (2017). Experiences with Multi-Modal Collaborative Virtual Laboratory (MMCVL). In *2017 ieee third international conference on multimedia big data (bigmm 2017)* (p. 376-383). (IEEE 3rd International Conference on Multimedia Big Data (BigMM), Laguna Hills, CA, APR 19-21, 2017) doi: 10.1109/BigMM.2017.62

DeVellis, R. F. (2003). *Scale development.* Sage Publications, Inc.

DFC Intelligence Research. (2021). *Global video game consumer segmentation* (Tech. Rep.). Retrieved from `https://www.dfcint.com/product/video-game-consumer-segmentation-2/`

Dholakiya, N. D., Ferjencik, M., Schofield, D., & Kubik, J. (2019). Virtual learning for safety, why not a smartphone? *Process Safety in Progress*, *38*(2). doi: 10.1002/prs.12005

Dholakiya, N. S., Kubík, J., Brozek, J., & Sotek, K. (2019). Unity3d game engine applied to chemical safety education. In *Lecture notes in electrical engineering* (pp. 81–87). Springer International Publishing. doi: 10.1007/978-3-030-21507-1_12

Dimitriadou, A., Djafarova, N., Turetken, O., Verkuyl, M., & Ferworn, A. (2020). Challenges in serious game design and development: Educators' experiences. *Simulation & Gaming*, *52*(2), 132–152. doi: 10.1177/1046878120944197

Domin, D. S. (1999). A review of laboratory instruction styles. *Journal of Chemical Education*, *76*(4), 543. doi: 10.1021/ed076p543

Donnelly, D., O'Reilly, J., & McGarr, O. (2013). Enhancing the Student Experiment Experience: Visible Scientific Inquiry Through a Virtual Chemistry Laboratory. *Research in Science Education*, *43*(4), 1571-1592. doi: 10.1007/s11165-012-9322-1

Duan, X., Kang, S.-J., Choi, J. I., & Kim, S. K. (2020). Mixed Reality System for Virtual Chemistry Lab. *KSII Transactions on Internet and Information Systems*, *14*(4), 1673-1688. doi: 10.3837/tiis.2020.04.014

Dunnagan, C. L., Dannenberg, D. A., Cuales, M. P., Earnest, A. D., Gurnsey, R. M., & Gallardo-Williams, M. T. (2020). Production and Evaluation of a Realistic Immersive Virtual Reality Organic Chemistry Laboratory Experience: Infrared Spectroscopy. *Journal of Chemical Education*, *97*(1), 258-262. doi: 10.1021/acs.jchemed.9b00705

Elston, H. J., & Luttrell, W. E. (1997). Designing effective laboratory safety training programs. *Chem. Health Saf.*, *4*(4), 11–44. doi: 10.1021/acs.chas.8b04407

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. doi: 10.1518/001872095779049543

Enneking, K. M., Breitenstein, G. R., Coleman, A. F., Reeves, J. H., Wang, Y., & Grove, N. P. (2019). The Evaluation of a Hybrid, General Chemistry Laboratory Curriculum: Impact on Students' Cognitive, Affective, and Psychomotor Learning. *Journal of Chemical Education*, *96*(6), 1058-1067. doi: 10.1021/acs.jchemed.8b00637

Faulconer, E. K., & Gruss, A. B. (2018). A review to weigh the pros and cons of online, remote, and distance science laboratory experiences. *The International Review of Research in Open and Distributed Learning*, *19*(2). doi: 10.19173/irrodl.v19i2.3386

Ferreira, C. (2019). *How to optimize your oculus quest app w/ renderdoc: Walkthroughs of key usage scenarios and optimization tips.* Retrieved 2022-03-15, from https://developer.oculus.com/blog/how-to-optimize-your-oculus-quest-app-w-renderdoc-walkthroughs-of-key-usage-scenarios-and-optimization-tips-part-1/

Ferrell, J. B., Campbell, J. P., McCarthy, D. R., McKay, K. T., Hensinger, M., Srinivasan, R., . . . Schneebeli, S. T. (2019). Chemical Exploration with Virtual Reality in Organic Teaching Laboratories. *JOURNAL OF CHEMICAL EDUCATION*, *96*(9), 1961-1966. doi: 10.1021/acs.jchemed.9b00036

Fivizzani, K. P. (2005). The evolution of chemical safety training. *Chemical Health and Safety*, *12*(6), 11–15. doi: 10.1016/j.chs.2005.08.005

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. doi: 10.1037/0003-066x.34.10.906

Fransson, G., Holmberg, J., & Westelius, C. (2020). The challenges of using head mounted virtual reality in k-12 schools from a teacher perspective. *Education and Information Technologies*, *25*(4), 3383–3404. doi: 10.1007/s10639-020-10119-1

Gal, Y. K., Uzan, O., Belford, R., Karabinos, M., & Yaron, D. (2015). Making Sense of Students' Actions in an Open-Ended Virtual Laboratory Environment. *Journal of Chemical Education*, *92*(4), 610-616. doi: 10.1021/ed500531a

Garcia Fracaro, S., Chan, P., Gallagher, T., Tehreem, Y., Toyoda, R., Bernaerts, K., . . . Wilk, M. (2021). Towards design guidelines for virtual reality training for the chemical industry. *Education for Chemical Engineers*, *36*, 12–23. doi: 10.1016/j.ece.2021.01.014

Garcia Fracaro, S., Glassey, J., Bernaerts, K., & Wilk, M. (2022). Immersive technologies for the training of operators in the process industry: A systematic literature review. *Computers & Chemical Engineering*, *160*, 107691. doi: 10.1016/j.compchemeng.2022.107691

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, *33*(4), 441–467.

Georgiadis, K., van Lankveld, G., Bahreini, K., & Westera, W. (2018, sep). Accommodating stealth assessment in serious games: Towards developing a generic tool. In *2018 10th international conference on virtual worlds and games for serious applications (VS-games).* IEEE. doi: 10.1109/vs-games.2018.8493409

Georgiou, J., Dimitropoulos, K., & Manitsaris, A. (2007). A virtual reality laboratory for distance education in chemistry. *International Journal of Educational and Pedagogical Sciences*, *1*(11), 617 - 624.

Gervasi, O., Riganelli, A., Pacifici, L., & Lagana, A. (2004). VMSLab-G: a virtual laboratory prototype for molecular science on the Grid. *Future Generation Computer Systems - The International Journal of E-Science*, *20*(5), 717-726. (International Conference on Computational Sciences, St Petersburg, RUSSIA, JUN, 2003) doi: 10.1016/j.future.2003.11.015

GIMP. (2022). *Gimp image editing software.* Retrieved 2022-03-15, from https://www.gimp.org/

Gloria, A. D., Bellotti, F., & Berta, R. (2014). Serious games for education and training. *International Journal of Serious Games*, *1*(1). doi: 10.17083/ijsg.v1i1.11

Gopalaswami, N., & Han, Z. (2020). Analysis of laboratory incident database. *Journal of Loss Prevention in the Process Industries*, *64*, 104027. doi: 10.1016/j.jlp.2019.104027

Grabowski, A., & Jankowski, J. (2015). Virtual reality-based pilot training for underground coal miners. *Safety Science*, *72*, 310–314. doi: 10.1016/j.ssci.2014.09.017

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology*, *5*(3), 347–358. doi: 10.1037/1076-8998.5.3.347

Gronbacher, G. (2004). *Pay attention! to safety. how to eliminate boredom in safety training.* Retrieved 2020-10-10, from https://www.thefabricator.com/thewelder/article/safety/pay-attention-to-safety

Han, J., Tian, Y., Song, W., & Fong, S. (2017). An Implementation of VR Chemistry Experiment System. In *International conference on big data and internet of things (bdiot 2017)* (p. 205-208). (International Conference on Big Data and Internet of Things (BDIOT), London, ENGLAND, DEC 20-22, 2017) doi: 10.1145/3175684.3175708

Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., & Vine, S. J. (2020, mar). A framework for the testing and validation of simulated environments in experimentation and training. *Frontiers in Psychology*, *11*. doi: 10.3389/fpsyg.2020.00605

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi: 10.3102/003465430298487

Hawkins, I., & Phelps, A. J. (2013). Virtual laboratory vs. traditional laboratory: which is more effective for teaching electrochemistry? *Chemistry Education Research and Practice*, *14*(4), 516-523. doi: 10.1039/c3rp00070b

Hedlund, A., Gummesson, K., Rydell, A., & Andersson, I.-M. (2016). Safety motivation at work: Evaluation of changes from six interventions. *Safety Science*, *82*, 155–163. doi: 10.1016/j.ssci.2015.09.006

Hensen, C., & Barbera, J. (2019). Assessing Affective Differences between a Virtual General Chemistry Experiment and a Similar Hands-On Experiment. *Journal of Chemical Education*, *96*(10), 2097-2108. doi: 10.1021/acs.jchemed.9b00561

Hensen, C., Glinowiecka-Cox, G., & Barbera, J. (2020). Assessing Differences between Three Virtual General Chemistry Experiments and Similar Hands-On Experiments. *Journal of Chemical Education*, *97*(3), 616-625. doi: 10.1021/acs.jchemed.9b00748

Herga, N. R., Cagran, B., & Dinevski, D. (2016). Virtual Laboratory in the Role of Dynamic Visualisation for Better Understanding of Chemistry in Primary School. *Eurasia Journal of Mathematics Science and Technology Education*, *12*(3), 593-608.

Herga, N. R., & Dinevski, D. (2012). Virtual Laboratory in Chemistry - Experimental Study of Understanding, Reproduction and Application of Acquired Knowledge of Subject's Chemical Content. *Organizacija*, *45*(3), 108-116. doi: 10.2478/v10051-012-0011-7

Herga, N. R., Glazar, S. A., & Dinevski, D. (2015). Dynamic visualization in the virtual laboratory enhances the fundamental understanding of chemical concepts. *Journal of Baltic Science Education*, *14*(3), 351-365.

Hew, K. F., Lan, M., Tang, Y., Jia, C., & Lo, C. K. (2019). Where is the "theory" within the field of educational technology research? *British Journal of Educational Technology*, *50*(3), 956–971. doi: 10.1111/bjet.12770

Hill, R. (2010). *Laboratory safety for chemistry students.* Hoboken, N.J: Wiley.

Hirzle, T., Cordts, M., Rukzio, E., Gugenheimer, J., & Bulling, A. (2021). A critical assessment of the use of ssq as a measure of general discomfort in vr head-mounted displays. In *Proceedings of the 2021 chi conference on human factors in computing systems.* New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3411764.3445361

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. doi: 10.1016/j.chb.2010.05.031

Howard, J., Gagné, M., Morin, A. J., & den Broeck, A. V. (2016). Motivation profiles at work: A self-determination theory approach. *Journal of Vocational Behavior*, *95-96*, 74–89. doi: 10.1016/j.jvb.2016.07.004

Hu, Y., Gallagher, T., Wouters, P., van der Schaaf, M., & Kester, L. (2022). Game-based learning has good chemistry with chemistry education: A three-level meta-analysis. *Journal of Research in Science Teaching*, *59*(9), 1499–1543. doi: 10.1002/tea.21765

Huang, W. D., & Johnson, T. (2009). Instructional game design using cognitive load theory. In *Handbook of research on effective electronic gaming in education* (pp. 1143–1165). IGI Global. doi: 10.4018/978-1-59904-808-6

.ch066

Ikhsan, J., Sugiyarto, K. H., & Astuti, T. N. (2020). Fostering student's critical thinking through a virtual reality laboratory. *International Journal of Interactive Mobile Technologies (iJIM)*, *14*(08), 183. doi: 10.3991/ijim.v14i08.13069

Ikram, W., Jeong, Y., Lee, B., Um, K., & Cho, K. (2015). Smart Virtual Lab Using Hand Gestures. In Park, JJ and Chao, HC and Arabnia, H and Yen, NY (Ed.), *Advanced Multimedia and Ubiquitous Engineering: Future Information Technology* (Vol. 352, p. 165-170). (10th International Scientific Conference on Future Information Technology (FutureTech), Hanoi, VIETNAM, MAY 18-20, 2015) doi: 10.1007/978-3-662-47487-7_25

Irby, S. M., Borda, E. J., & Haupt, J. (2018). Effects of Implementing a Hybrid Wet Lab and Online Module Lab Curriculum into a General Chemistry Course: Impacts on Student Performance and Engagement with the Chemistry Triplet. *Journal of Chemical Education*, *95*(2), 224-232. doi: 10.1021/acs.jchemed.7b00642

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

Jagodzinski, P., & Wolski, R. (2014). The examination of the impact on students' use of gestures while working in a virtual chemical laboratory for their cognitive abilities. *Problems of Education in the 21st Century*, *61*, 46-57.

Jagodzinski, P., & Wolski, R. (2015). Assessment of Application Technology of Natural User Interfaces in the Creation of a Virtual Chemical Laboratory. *Journal of Science Education and Technology*, *24*(1), 16-28. doi: 10.1007/s10956-014-9517-5

Johnson, C. I., Bailey, S. K. T., & Buskirk, W. L. V. (2017, nov). Designing effective feedback messages in serious games and simulations: A research review. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 119–140). Springer International Publishing. doi: 10.1007/978-3-319-39298-1_7

Johnstone, A. H. (1991). Why is science difficult to learn? things are seldom what they seem. *Journal of Computer Assisted Learning*, *7*(2), 75–83. doi: 10.1111/j.1365-2729.1991.tb00230.x

Jones, N. (2018). Simulated labs are booming. *Nature*, *562*(7725), S5–S7. doi: 10.1038/d41586-018-06831-1

Jordá, J. M. M. (2013). Virtual tools: Virtual laboratories for experimental science – an experience with VCL tool. *Procedia - Social and Behavioral Sciences*, *106*, 3355–3365. doi: 10.1016/j.sbspro.2013.12.388

Kalyuga, S., & Plass, J. L. (2009). Evaluating and managing cognitive load in games. In *Handbook of research on effective electronic gaming in education* (pp. 719–737). IGI Global. doi: 10.4018/978-1-59904-808-6.ch041

Kastrenakes, J., & Heath, A. (2021). *Facebook is spending at least $10 billion this year on its metaverse division.* Retrieved from https://www.theverge.com/2021/10/25/22745381/facebook-reality-labs-10-billion-metaverse

Kebritchi, M., & Hirumi, A. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers & Education*, *51*(4), 1729 - 1743. doi: 10.1016/j.compedu.2008.05.004

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203-220. doi: 10.1207/s15327108ijap0303_3

Kerby, D. S. (2014). The simple difference formula: An approach to teaching nonparametric correlation. *Comprehensive Psychology*, *3*. doi: 10.2466/11.it.3.1

Kim, H., Nah, S., Oh, J., & Ryu, H. (2019). Vr-moocs: A learning management system for vr education. In *2019 26th ieee conference on virtual reality and 3d user interfaces (vr)* (p. 1325-1326). (26th IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, JAPAN, MAR 23-27, 2019)

Kim, Y., Cho, S., Fong, S., Park, Y. W., & Cho, K. (2017). Design of hand gesture interaction framework on clouds for multiple users. *Journal of Supercomputing*, *73*(7, SI), 2851-2866. doi: 10.1007/s11227-016-1722-y

Kim, Y., Xi, Y., Cho, S., Fong, S., Yi, C., Um, K., & Cho, K. (2016). Interaction Engine Design for Virtual Experiments by Multi-users. In Park, JJ and Jin, H and Jeong, YS and Khan, MK (Ed.), *Advanced multimedia and ubiquitous engineering futuretech & mue* (Vol. 393, p. 229-234). (11th International Scientific Conference on Future Information Technology (FutureTech) / 10th International Conference on Multimedia and Ubiquitous Engineering (MUE), Beijing, PEOPLES R CHINA, APR 20-22, 2016) doi: 10.1007/978-981-10-1536-6_31

Kirschner, P., & van Merriënboer, J. J. G. (2018, 01). Ten steps to complex learning: a systematic approach to four-component instructional design (3rd ed.), by jeroen j. g. van merriënboer and paul a. kirschner. *TechTrends*, *62*(2), 204–205. doi: 10.1007/s11528-018-0254-0

Kolb, D. (1984). *Experiential learning : experience as the source of learning and development.* Englewood Cliffs, N.J: Prentice-Hall.

Kolil, V. K., Muthupalani, S., & Achuthan, K. (2020). Virtual experimental platforms in chemistry laboratory education and its impact on experimental self-efficacy. *International Journal of Education Technology in Higher Education*, *17*(1). doi: 10.1186/s41239-020-00204-3

Kourtesis, P., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2019). Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: A technological review and meta-analysis. *Frontiers in Human Neuroscience*, *13*. doi: 10.3389/fnhum.2019.00342

Laberge, M., MacEachen, E., & Calvet, B. (2014). Why are occupational health and safety training approaches not effective? understanding young worker learning processes using an ergonomic lens. *Safety Science*, *68*, 250–257. doi: 10.1016/j.ssci.2014.04.012

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? a structural equation modeling approach. *Computers & Education*, *55*(4), 1424–1442. doi: 10.1016/j.compedu.2010.06.006

Li, X., Yi, W., Chi, H.-L., Wang, X., & Chan, A. P. (2018). A critical review of virtual and augmented reality (vr/ar) applications in construction safety. *Automation in Construction*, *86*, 150 - 162. doi: 10.1016/j.autcon.2017.11.003

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 227–246). Cambridge University Press. doi: 10.1017/cbo9781139547369.012

Lynch, T., & Ghergulescu, I. (2017). Review of virtual labs as the emerging technologies for teaching stem subjects. In *INTED2017 proceedings.* IATED. doi: 10.21125/inted.2017.1422

Ma, J., & Nickerson, J. V. (2006). Hands-on, simulated, and remote laboratories: A comparative literature review. *ACM Comput. Surv.*, *38*(3), 7–es. doi: 10.1145/1132960.1132961

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2013). cluster: Cluster analysis basics and extensions [Computer software manual]. (R package version 1.14.4 — For new features, see the 'Changelog' file (in the package source))

Magyaródi, T., Nagy, H., Soltész, P., Mózes, T., & Oláh, A. (2013). Psychometric properties of a newly established flow state questionnaire. *The Journal of Happiness & Well-Being*, *1*(2), 85–96.

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, *35*(6), 691–707. doi: 10.1111/jcal.12375

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141–1164. doi: 10.1007/s11423-018-9581-2

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, *33*(3), 937–958. doi: 10.1007/s10648-020-09586-2

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

Makransky, G., Wismer, P., & Mayer, R. E. (2019, JUN). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, *35*(3), 349-358. doi: {10.1111/jcal.12335}

Mallam, S. C., Nazir, S., Renganayagalu, S. K., Ernstsen, J., Veie, S., & Edwinson, A. E. (2019). Design of experiment comparing users of virtual reality head-mounted displays and desktop computers. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th congress of the international ergonomics association (iea 2018)* (pp. 240–249). Cham: Springer International Publishing.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science*, *5*(4), 333–369. doi: 10.1207/s15516709cog0504_2

Mariani, M. G., Petruzziello, G., Vignoli, M., & Guglielmi, D. (2022). Development and initial validation of the safety training engagement scale (STE-s). *European Journal of Investigation in Health, Psychology and Education*, *12*(8), 975–988. doi: 10.3390/ejihpe12080070

Martha, A. S. D., & Santoso, H. (2019). The design and impact of the pedagogical agent: A systematic literature review. *The Journal of Educators Online*, *16*(1). doi: 10.9743/jeo.2019.16.1.8

Martinez-Jimenez, P., Pontes-Pedrajas, A., Polo, J., & Climent-Bellido, M. (2003). Learning in chemistry with virtual laboratories. *Journal of Chemical Education*, *80*(3), 346-352.

Mayer, R. E. (2014a). *The cambridge handbook of multimedia learning* (R. Mayer, Ed.). Cambridge University Press. doi: 10.1017/cbo9781139547369

Mayer, R. E. (2014b). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 345–368). Cambridge University Press. doi: 10.1017/cbo9781139547369.017

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 279–315). Cambridge University Press. doi: 10.1017/cbo9781139547369.015

Ménard, A. D., & Trant, J. F. (2019). A review and critique of academic lab safety research. *Nature Chemistry*, *12*(1), 17–25. doi: 10.1038/s41557-019-0375-x

Merck Group. (2021). *From video game to business case.* Retrieved from `https://www.merckgroup.com/en/the-future-transformation/virtual-augmented-reality.html`

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

Miliszewska, I., & Sztendur, E. (2011). Playing it safe: Approaching science safety awareness through computer game-based training. *Issues in Informing Science and Information Technology*, *8*, 037–047. doi: 10.28945/1403

Minocha, S., Tudor, A.-D., & Tilling, S. (2017). Affordances of mobile virtual reality and their role in learning and teaching. BCS Learning & Development. doi: 10.14236/ewic/hci2017.44

Moher, D., Liberati, A., Tetzlaff, J., & and, D. G. A. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, *339*(jul21 1), b2535–b2535. doi: 10.1136/bmj.b2535

Moozeh, K., Farmer, J., Tihanyi, D., & Evans, G. J. (2020). Learning Beyond the Laboratory: A Web Application Framework for Development of Interactive Postlaboratory Exercises. *Journal of Chemical Education*, *97*(5), 1481-1486. doi: 10.1021/acs.jchemed.9b00756

Moran, C. M., Diefendorff, J. M., Kim, T.-Y., & Liu, Z.-Q. (2012). A profile approach to self-determination theory motivations at work. *Journal of Vocational Behavior*, *81*(3), 354–363. doi: 10.1016/j.jvb.2012.09.002

Morozov, M., Tanakov, A., Gerasimov, A., Bystrov, D., & Cvirko, V. (2004). Virtual chemistry laboratory for school education. In Cunningham, P and Cunningham, M (Ed.), *Proceedings of the ieee international conference on advanced learning technologies (icalt'04)* (Vol. 1, p. 1742-1748). IEEE. (eChallenges e-2004 Conference, Vienna, AUSTRIA, OCT 27-29, 2004) doi: 10.1109/ICALT.2004.1357486

Nicholson, S. (2015). A recipe for meaningful gamification. In *Gamification in education and business.*

Novet, J. (2021). Microsoft wins u.s. army contract for augmented reality headsets, worth up to $21.9 billion over 10 years. *CNBC*. Retrieved from `https://www.cnbc.com/2021/03/31/microsoft-wins-contract-to-make-modified-hololens-for-us-army.html`

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955. doi: 10.1002/asi.20801

Owen, M. (2022). *Apple reportedly working on at least three ar, vr headsets.*

Retrieved from https://www.theverge.com/2021/10/25/22745381/facebook-reality-labs-10-billion-metaverse

Papadopoli, R., Nobile, C. G. A., Trovato, A., Pileggi, C., & Pavia, M. (2020). Chemical risk and safety awareness, perception, and practices among research laboratories workers in italy. *Journal of Occupational Medicine and Toxicology*, *15*(1). doi: 10.1186/s12995-020-00268-x

Patle, D. S., Manca, D., Nazir, S., & Sharma, S. (2018). Operator training simulators in virtual reality environment for process operators: a review. *Virtual Reality*, *23*(3), 293–311. doi: 10.1007/s10055-018-0354-3

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, *13*(3), 423–451. doi: 10.1207/s15327809jls1303_6

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In *Motivational psychology of human development - developing motivation and motivating development* (pp. 143–163). Elsevier. doi: 10.1016/s0166-4115(00)80010-2

Penn, M., & Ramnarain, U. (2019). South African university students' attitudes towards chemistry learning in a virtually simulated learning environment. *Chemistry Education Research and Practice*, *20*(4), 699-709. doi: 10.1039/c9rp00014c

Petrock, V. (2021). *The pandemic pushed xr use beyond fun and games.* Retrieved from https://www.insiderintelligence.com/content/pandemic-pushed-xr-use-beyond-gaming

Pohl, D., & de Tejada Quemada, C. F. (2016). See what i see: Concepts to improve the social acceptance of HMDs. In *2016 IEEE virtual reality (VR).* IEEE. doi: 10.1109/vr.2016.7504756

Prensky, M. (2001a). *Digital game-based learning.* New York: McGraw-Hill.

Prensky, M. (2001b). Digital natives, digital immigrants part 2: Do they really think differently? *On the Horizon*, *9*(6), 1–6. doi: 10.1108/10748120110424843

Pyatt, K., & Sims, R. (2012). Virtual and Physical Experimentation in Inquiry-Based Science Labs: Attitudes, Performance and Access. *Journal of Science Education and Technology*, *21*(1), 133-147. doi: 10.1007/s10956-011-9291-6

Qvist, P., Kangasniemi, T., Palomaki, S., Seppanen, J., Joensuu, P., Natri, O., . . . Nordstrom, K. (2015). Design of Virtual Learning Environments Learning Analytics and Identification of Affordances and Barriers. *International Journal of Engineering Pedagogy*, *5*(4), 64-75. doi: 10.3991/ijep.v5i4.4962

Ramadan, R., & Widyani, Y. (2013). Game development life cycle guidelines. In *2013 international conference on advanced computer science and information systems (ICACSIS).* IEEE. doi: 10.1109/icacsis.2013

.6761558

Ramos, S., Pimentel, E. P., d. G. B. Marietto, M., & Botelho, W. T. (2016). Hands-on and virtual laboratories to undergraduate chemistry education: Toward a pedagogical integration. In *2016 ieee frontiers in education conference (fie)* (p. 1-8). doi: 10.1109/FIE.2016.7757580

Rapeepisarn, K., Wong, K. W., Fung, C. C., & Khine, M. S. (2008). The relationship between game genres, learning techniques and learning styles in educational computer games. In Z. Pan, X. Zhang, A. El Rhalibi, W. Woo, & Y. Li (Eds.), *Technologies for e-learning and digital entertainment* (pp. 497–508). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rashid, T., & Asghar, H. M. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers in Human Behavior*, *63*, 604–612. doi: 10.1016/j.chb.2016.05.084

Ratamun, M. M., & Osman, K. (2018). The effectiveness of virtual lab compared to physical lab in the mastery of science process skills for chemistry experiment. *Problems of Education in the 21st Century*, *76*(4), 544-560.

Reid, N., & Shah, I. (2007). The role of laboratory work in university chemistry. *Chemistry Education Research and Practice*, *8*, 172-185. doi: 10.1039/B5RP90026C

Report Linker. (2022). *Extended reality (xr) market - growth, trends, covid-19 impact, and forecasts (2021 - 2026).* Retrieved from `https://www.globenewswire.com/news-release/2022/06/15/2462976/0/en/Extended-Reality-XR-Market-Growth-Trends-COVID-19-Impact-and-Forecasts-2021-2026.html`

Romero-Rondón, M. F., Sassatelli, L., Precioso, F., & Aparicio-Pardo, R. (2018, jun). Foveated streaming of virtual reality videos. In *Proceedings of the 9th ACM multimedia systems conference.* ACM. doi: 10.1145/3204949.3208114

Rowe, R. J., Koban, L., Davidoff, A. J., & Thompson, K. H. (2018). Efficacy of Online Laboratory Science Courses. *Journal of Formative Design in Learning*, *2*(1), 56-67. doi: 10.1007/s41686-017-0014-0

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. doi: 10.1037/0003-066x.55.1.68

Ryan, R. M., & Rigby, C. S. (2020). Motivational foundations of game-based learning. In J. L. P. R. E. Mayer & B. D. Homer (Eds.), *Handbook of game-based learning* (p. 153-177). The MIT Press.

Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. *Construction Management and Economics*, *31*(9), 1005–1017. doi: 10.1080/01446193.2013.828844

Salen, K., & Zimmerman, E. (2004). *Rules of play.* MIT Press Ltd. Retrieved from `https://www.ebook.de/de/product/3675147/katie_salen_professor_tekinbas_eric_zimmerman_rules_of_play.html`

Sampaio, P., Mendonca, R., & Carreira, S. (2014). Learning chemistry with VirtualLabs@Uma: a customizable 3D platform for new experimental protocols. *Multimedia Tools and Applications*, *71*(3), 1129-1155. doi: 10.1007/s11042-012-1260-4

Scherer, R., & Tiemann, R. (2012). Factors of problem-solving competency in a virtual chemistry environment: The role of metacognitive knowledge about strategies. *Computers & Education*, *59*(4), 1199–1214. doi: 10.1016/j.compedu.2012.05.020

Schrier, K. (2014). *Learning, education and games volume one: Curricular and design considerations* (K. Shrier, Ed.). ETC Press. Retrieved from `https://www.ebook.de/de/product/23974280/et_al_karen_schrier_learning_education_and_games.html`

Schröder, I., Huang, D. Y. Q., Ellis, O., Gibson, J. H., & Wayne, N. L. (2016). Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers. *Journal of Chemical Health and Safety*, *23*(1), 12 - 23. doi: 10.1016/j.jchas.2015.03.001

Scott, N. (2016). *Enjoyment, values, pressure, or something else: What influences employees' safety behaviours?* (Unpublished doctoral dissertation). Saint Mary's University.

Scott, N., Fleming, M., & Kelloway, E. K. (2014). Understanding why employees behave safely from a self-determination theory perspective. In *The oxford handbook of work engagement, motivation, and self-determination theory.* Oxford University Press. doi: 10.1093/oxfordhb/9780199794911.013.019

Seery, M. K. (2020). Establishing the laboratory as the place to learn how to do chemistry. *Journal of Chemical Education*, *97*(6), 1511–1514. doi: 10.1021/acs.jchemed.9b00764

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer games and instruction*, *55*(2), 503–524.

Simmons, H. E., Matos, B., & Simpson, S. A. (2017). Analysis of injury data to improve safety and training. *J. Chem. Health Saf.*, *24*(1), 21–28. doi: 10.1021/acs.chas.8b24107

Sketchfab. (2022). *Sketchfab online platform for 3d models.* Retrieved 2022-03-15, from `https://sketchfab.com/`

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. doi: 10.1162/pres.1997.6.6.603

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

Solikhin, F., Sugiyarto, K. H., & Ikhsan, J. (2019). The impact of virtual laboratory integrated into hybrid learning use on students' achievement. *Jurnal Ilmiah Peuradeun*, *7*(1), 81-94. doi: 10.26811/peuradeun.v7i1.268

Somrak, A., Humar, I., Hossain, M. S., Alhamid, M. F., Hossain, M. A., & Guna, J. (2019, may). Estimating VR sickness and user experience using different HMD technologies: An evaluation study. *Future Generation Computer Systems*, *94*, 302–316. doi: 10.1016/j.future.2018.11.041

Srinivasan, B., Iqbal, M. U., Shahab, M. A., & Srinivasan, R. (2022). Review of virtual reality (VR) applications to enhance chemical safety: From students to plant operators. *ACS Chemical Health & Safety*, *29*(3), 246–262. doi: 10.1021/acs.chas.2c00006

Stack Overflow. (2022). *Stack overflow community forum.* Retrieved 2022-03-15, from https://stackoverflow.com/

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73–93. doi: 10.1111/j.1460-2466.1992.tb00812.x

Su, C.-H., & Cheng, T.-W. (2019). A Sustainability Innovation Experiential Learning Model for Virtual Reality Chemistry Laboratory: An Empirical Study with PLS-SEM and IPMA. *Sustainability*, *11*(4). doi: 10.3390/su11041027

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296. doi: 10.1023/a:1022193728205

Sypsas, A., & Kalles, D. (2018). Virtual Laboratories in Biology, Biotechnology and Chemistry education: A Literature Review. In Karanikolas, NN and Mamalis, B (Ed.), *22nd pan-hellenic conference on informatics (pci2018)* (p. 70-75). (22nd Pan-Hellenic Conference on Informatics (PCI), Univ W Attica, Dept Informat & Comp Engn, Athens, GREECE, NOV 29-DEC 01, 2018) doi: 10.1145/3291533.3291560

Tarng, W., Hsie, C.-C., Lin, C.-M., & Lee, C.-Y. (2017). Development and Application of a Virtual Laboratory for Synthesizing and Analyzing Nanogold Particles. *Journal of Computers*, *12*(3), 270-283. doi: 10.17706/jcp.12.3.270-283

Tatli, Z., & Ayas, A. (2010). Virtual laboratory applications in chemistry education. In Keser, H and Ozcinar, Z and Kanbul, S (Ed.), *World conference on learning, teraaching and administration papers* (Vol. 9). (1st World Conference on Learning, Teaching and Administration (WCLTA), Amer Univ, Cairo, EGYPT, OCT 29-31, 2010) doi: 10.1016/j.sbspro.2010.12.263

Tatli, Z., & Ayas, A. (2012). Virtual chemistry laboratory: effect of constructivist learning environment. *Turkish Online Journal of Distance Education*, *13*(1), 183-199.

Tatli, Z., & Ayas, A. (2013). Effect of a Virtual Chemistry Laboratory on Students' Achievement. *Educational Technology & Society*, *16*(1), 159-170.

Trindade, J., Fiolhais, C., & Almeida, L. (2002). Science learning in virtual environments: a descriptive study. *British Journal of Educational Technology*, *33*(4), 471–488. doi: 10.1111/1467-8535.00283

Udeozor, C., Abegão, F., & Glassey, J. (2021, sep). Exploring log data for behaviour and solution pattern analyses in a serious game. In *Gamification and social networks in education* (pp. 209–227). MacroWorld. doi: 10.15340/978-625-00-0106-6_10

Ullah, S., Ali, N., & Rahman, S. U. (2016). The Effect of Procedural Guidance on Students' Skill Enhancement in a Virtual Chemistry Laboratory. *Journal of Chemical Education*, *93*(12), 2018-2025. doi: 10.1021/acs.jchemed.5b00969

Unity Answers. (2022). *Unity answers community forum.* Retrieved 2022-03-15, from `https://answers.unity.com/index.html`

Unity Asset Store. (2022). *Unity asset store.* Retrieved 2022-03-15, from `https://assetstore.unity.com/`

Unity Documentation. (2022). *Unity documentation.* Retrieved 2022-03-15, from `https://docs.unity.com/`

Unity Learn. (2022). *Unity learn tutorial platform.* Retrieved 2022-03-15, from `https://learn.unity.com/`

Unity Youtube. (2022). *Unity's youtube channel.* Retrieved 2022-03-15, from `https://www.youtube.com/c/unity`

Unity3D. (2022). *Unity3d game engine.* Retrieved 2022-03-15, from `https://unity.com/`

Valem. (2022). *Valem's youtube channel.* Retrieved 2022-03-15, from `https://www.youtube.com/c/valemvr`

Van den Broeck, A., Lens, W., De Witte, H., & Coillie, H. V. (2013). Unraveling the importance of the quantity and the quality of workers' motivation for well-being: A person-centered perspective. *Journal of Vocational Behavior*, *82*(1), 69–78. doi: 10.1016/j.jvb.2012.11.005

van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 263–278). Cambridge University Press. doi: 10.1017/cbo9781139547369.014

van Merriënboer, J. J. G., & Kester, L. (2014). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 104–148). Cambridge University Press. doi: 10.1017/cbo9781139547369

.007

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology*, *101*(3), 671–688. doi: 10.1037/a0015083

van Wyk, E., & de Villiers, M. R. R. (2014). Applying design-based research for developing virtual reality training in the south african mining industry. In *Proceedings of the southern african institute for computer scientist and information technologists annual conference 2014 on saicsit 2014 empowered by technology* (pp. 70:70–70:81). New York, NY, USA: ACM. doi: 10.1145/2664591.2664627

Villiers, M. R. R. D. (2012). Models for interpretive information systems research, part 2. In *Research methodologies, innovations and philosophies in software systems engineering and information systems* (pp. 238–255). IGI Global. doi: 10.4018/978-1-4666-0179-6.ch012

Vinodkumar, M., & Bhasi, M. (2010). Safety management practices and safety behaviour: Assessing the mediating role of safety knowledge and motivation. *Accident Analysis & Prevention*, *42*(6), 2082–2093. doi: 10.1016/j.aap.2010.06.021

VR with Andrew. (2022). *Vr with andrew's youtube channel.* Retrieved 2022-03-15, from https://www.youtube.com/c/VRwithAndrew

Walsh, K. R., & Pawlowski, S. D. (2002). Virtual reality: A technology in need of IS research. *Communications of the Association for Information Systems*, *8*. doi: 10.17705/1cais.00820

Walters, A. U., Lawrence, W., & Jalsa, N. K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students. *Safety Science*, *96*, 161–171. doi: 10.1016/j.ssci.2017.03.017

Wang, A., & Dong, A. (2018). Design of Virtual Chemical Experiment Platform Based on Unity3D. In Tseng, J and Kotenko, I (Ed.), *3rd annual international conference on information system and artificial intelligence (isai2018)* (Vol. 1069). (3rd Annual International Conference on Information System and Artificial Intelligence (ISAI), Suzhou, PEOPLES R CHINA, JUN 22-24, 2018) doi: 10.1088/1742-6596/1069/1/012133

Ward, L. (2021). *Market analysis perspective: Worldwide gaming and esports* (Tech. Rep.). Retrieved from https://www.idc.com/getdoc.jsp?containerId=US47267721

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, *10*. doi: 10.3389/fpsyg.2019.00158

Weitze, C. (2014). Developing goals and objectives for gameplay and learning. In K. Schrier (Ed.), *Learning, education and games volume one: Curricular and design considerations* (p. 225-249). ETC Press.

Wenger, J. L. E. (1991). *Situated learning.* Cambridge University Press. Retrieved from `https://www.ebook.de/de/product/2985304/jean_lave_etienne_wenger_situated_learning.html`

Winkelmann, K., Keeney-Kennicutt, W., Fowler, D., & Macik, M. (2017). Development, Implementation, and Assessment of General Chemistry Lab Experiments Performed in the Virtual World of Second Life. *Journal of Chemical Education*, *94*(7), 849-858. doi: 10.1021/acs.jchemed.6b00733

Winkelmann, K., Keeney-Kennicutt, W., Fowler, D., Macik, M. L., Guarda, P. P., & Ahlborn, C. J. (2020). Learning gains and attitudes of students performing chemistry experiments in an immersive virtual world. *Interactive Learning Environments*, *28*(5, SI), 620-634. doi: 10.1080/10494820.2019.1696844

Winkelmann, K., Scott, M., & Wong, D. (2014). A Study of High School Students' Performance of a Chemistry Experiment within the Virtual World of Second Life. *Journal of Chemical Education*, *91*(9, SI), 1432-1438. doi: 10.1021/ed500009e

Withers, J. H., Freeman, S. A., & Kim, E. (2012). Learning and retention of chemical safety training information: A comparison of classroom versus computer-based formats on a college campus. *Journal of Chemical Health and Safety*, *19*(5), 47–55. doi: 10.1016/j.jchas.2011.12.001

Wolski, R., & Jagodzinski, P. (2019). Virtual laboratory-Using a hand movement recognition system to improve the quality of chemical education. *British Journal of Educational Technology*, *50*(1), 218-231. doi: 10.1111/bjet.12563

Woodfield, B. F., Andrus, M. B., Andersen, T., Miller, J., Simmons, B., Stanger, R., … Bodily, G. (2005). The virtual ChemLab project: A realistic and sophisticated simulation of organic synthesis and organic qualitative analysis. *Journal of Chemical Education*, *82*(11), 1728. doi: 10.1021/ed082p1728

Woodfield, B. F., Catlin, H. R., Waddoups, G. L., Moore, M. S., Swan, R., Allen, R., & Bodily, G. (2004). The virtual ChemLab project: A realistic and sophisticated simulation of inorganic qualitative analysis. *Journal of Chemical Education*, *81*(11), 1672. doi: 10.1021/ed081p1672

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. doi: 10.1037/a0031311

Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, *60*(1), 412–425. doi: 10.1016/j.compedu.2012.07.018

Wu, B.-J., Wong, S.-K., & Li, T.-W. (2019). Virtual titration laboratory experiment with differentiated instruction. *Computer Animation and*

Virtual Worlds, *30*(3-4). (Conference on Computer Animation and Social Agents (CASA), Paris, FRANCE, 2019) doi: 10.1002/cav.1882

Yaron, D., Karabinos, M., Lange, D., Greeno, J. G., & Leinhardt, G. (2010). The ChemCollective-Virtual Labs for Introductory Chemistry Courses. *Science*, *328*(5978), 584-585. doi: 10.1126/science.1182435

Zayas-Pérez, B., & Cox, R. (2009). Teaching safety precautions in a laboratory dve: the effects of information location and interactivity. *Computación y Sistemas*, *13*.

Zhong, Y., & Liu, C. (2014). A domain-oriented end-user design environment for generating interactive 3D virtual chemistry experiments. *Multimedia Tools and Applications*, *72*(3), 2895-2924. doi: 10.1007/s11042-013-1554-1

Biography

Philippe Chan is a chemist/chemical engineer from Belgium. He grew up in Blankenberge where he also received his secondary school diploma in Sciences-Mathematics in 2012 at the Sint-Pieters/Sint-Jozefshandelschool. He then completed his Master of Science in Chemistry degree in 2018 at Ghent University. After his university degree, he pursued his next ambition to earn a Doctor of Philosophy (PhD) degree.

In 2019, he was awarded a Marie Skłodowska-Curie Actions scholarship for the ETN-CHARMING project. During his doctoral research, he investigated the implementation of a VR serious game as a training tool for chemical lab safety to increase engagement and motivation of trainees in the chemical industry and academia. For this research, he designed, developed and evaluated a VR serious game, called VR LaboSafe Game. Despite being a chemist, he was still eager to learn interdisciplinary skills in the fields of game design, game development, instructional and motivational psychology. His work was mainly conducted at the chemical company Arkema in Lyon, France. Then, he carried out the last phase of his doctoral research in KU Leuven, Belgium.

During his studies and research, he developed a variety of technical skills: game development with Unity3D and C# language; 3D modelling with Blender; visual editing with Adobe Photoshop, Illustrator and Premiere Pro; statistical programming with R language; and knowledge about process and occupational safety. Also soft skills were improved: presentation skills; international communication; and team collaboration.

From early in his life, he always has been curious about understanding how the world works at a molecular level, how to use technology to simplify procedures and how to help other people to get motivated. This has led to a never-ending passion for chemistry, a lively enthusiasm for digital technology and a growing interest in motivational psychology. His constant search for the next challenging goal has brought him accomplishments in earning a master and now a doctoral degree. By combining his creative and interdisciplinary mindset, his next ambition is to become an entrepreneur, passionate to support the chemical industry in the implementation of VR/AR solutions.

He can speak fluently four languages, including Dutch (native), Cantonese (native), English (very high proficiency) and French (professional proficiency). Also in his free time he likes to physically challenge himself by doing rock climbing, parkour and skateboarding. Otherwise, he likes to relax by cooking and playing games.

LinkedIn: `https://linkedin.com/in/philippe-chan`
ORCiD: `https://orcid.org/0000-0003-2146-611X`

List of publications

International journals with review

- Garcia Fracaro S. , **Chan P.**, Gallagher T., Tehreem Y., Toyoda R., Bernaerts K., Glassey J., Pfeiffer T., Slof B., Wachsmuth S., Wilk M. (2021). Towards design guidelines for virtual reality training for the chemical industry. *Education for Chemical Engineers*, Elsevier BV, 36, 12-23. `https://doi.org/10.1016/j.ece.2021.01.014`
- **Chan P.**, Van Gerven T., Dubois J.-L., Bernaerts K. (2021). Virtual chemical laboratories: a systematic literature review of research, technologies and instructional design. *Computers & Education Open*, Volume 2. `https://doi.org/10.1016/j.caeo.2021.100053`
- **Chan P.**, Van Gerven T., Dubois J.-L., Bernaerts K. Study of motivation and engagement for chemical laboratory safety training with VR serious game. *Safety Science.* (in review March 2023)

International conference proceedings with review - full papers

- **Chan P.**, Van Gerven T., Dubois J.-L., Bernaerts K. (2021). Design and Development of a VR Serious Game for Chemical Laboratory Safety. In: de Rosa F., Marfisi Schottman I., Baalsrud Hauge J., Bellotti F., Dondio P., Romero M. (eds) Games and Learning Alliance GALA 2021. *Lecture Notes in Computer Science*, vol 13134. Springer, Cham. `https://doi.org/10.1007/978-3-030-92182-8_3`

International conference proceedings with review - abstracts

- **Chan P.**, Guilpain G., Dubois J.-L., Toyoda R., Abegão F. R., Glassey J. , Gallagher T. , Kester L. , Bernaerts K., Van Gerven T. (2021). VR LaboSafe Game: a virtual reality training game for chemical laboratory safety. 13th European Congress of Chemical Engineering and 6th European Congress of Applied Biotechnology (ECCE-ECAB2021). Online

Local conference proceedings with review - abstracts

- **Chan P.**, Guilpain G., Dubois J.-L., Bernaerts K., Van Gerven T. (2021). VR LaboSafe Game: un jeu vidéo en réalité virtuelle pour la formation en sécurité dans les laboratoires chimiques. JT Couleur et Evolution de l'enseignement. Lyon, France. https://new.societechimiquedefrance.fr/divisions/enseignement-formation/mercredi-30-juin-2021/

Outreach contributions

Outreach title	**Type**	**Date**	**Link**
1. Educational Game on Recycling and Sustainability	Blog article	11/08/2019	https://charming-etn.eu/2019/08/11/educational-game-on-recycling-and-sustainability/
2. Advertisement for CHARMING Summer School	LinkedIn post	08/09/2019	https://www.linkedin.com/feed/update/urn:li:activity:6575018831238971394/
3. Arkema Centre de Recherche Rhône-Alpes (CRRA) – Innovative Chemical Research	Blog article	30/07/2020	https://charming-etn.eu/2020/07/30/arkema-centre-de-recherche-rhone-alpes-crra-innovative-chemical-research/
4. VR LaboSafe Game demo version and video trailer	Demo app + Video	19/11/2021	demo version: https://github.com/PhilippeChan/VRLaboSafeGameDemo trailer: https://www.youtube.com/watch?v=yDEyEgvxA88
5. CHARMING in Lyon: 6th Network-Wide Event	Blog article	08/12/2021	https://charming-etn.eu/2021/12/08/charming-in-lyon-6th-network-wide-event/

Outreach title	**Type**	**Date**	**Link**
6. Episode 13: VR LaboSafe Game – A serious game for chemical labs	Podcast episode	01/03/2022	https://open.spotify.com/episode/7c7R3oEYYQSj9lIaX6wUyM
7. Immersive Chemistry Afternoon	Seminar	09/06/2022	https://www.youtube.com/watch?v=oitKxJHVzbs
8. ETN-CHARMING Knowledge clips: ESR11 Philippe Chan	Video	26/12/2022	https://www.youtube.com/watch?v=6LUw4sKmmr8

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF CHEMICAL ENGINEERING
BIO- AND CHEMICAL SYSTEMS TECHNOLOGY, REACTOR ENGINEERING AND SAFETY (CREAS)
Celestijnenlaan 200F box 2424
B-3001 Leuven
philippe.chan@kuleuven.be
https://cit.kuleuven.be/creas