



# An immersive virtual reality learning environment with CFD simulations: Unveiling the Virtual Garage concept

Serkan Solmaz<sup>1</sup> · Liesbeth Kester<sup>2</sup> · Tom Van Gerven<sup>1</sup>

Received: 14 November 2022 / Accepted: 20 March 2023 / Published online: 3 April 2023  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Virtual reality has become a significant asset to diversify the existing toolkit supporting engineering education and training. The cognitive and behavioral advantages of virtual reality (VR) can help lecturers reduce entry barriers to concepts that students struggle with. Computational fluid dynamics (CFD) simulations are imperative tools intensively utilized in the design and analysis of chemical engineering problems. Although CFD simulation tools can be directly applied in engineering education, they bring several challenges in the implementation and operation for both students and lecturers. In this study, we develop the “Virtual Garage” as a task-centered educational VR application with CFD simulations to tackle these challenges. The Virtual Garage is composed of a holistic immersive virtual reality experience to educate students with a real-life engineering problem solved by CFD simulation data. The prototype is tested by graduate students ( $n = 24$ ) assessing usability, user experience, task load and simulator sickness via standardized questionnaires together with self-reported metrics and a semi-structured interview. Results show that the Virtual Garage is well-received by participants. We identify features that can further leverage the quality of the VR experience with CFD simulations. Implications are incorporated throughout the study to provide practical guidance for developers and practitioners.

**Keywords** Virtual reality · Computational fluid dynamics · Immersive learning · Engineering education

---

✉ Tom Van Gerven  
tom.vangerven@kuleuven.be

<sup>1</sup> Department of Chemical Engineering, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium

<sup>2</sup> Department of Education, Utrecht University, Utrecht, The Netherlands

## 1 Introduction

Applied engineering deals with complex problems to provide efficient solutions for a sustainable world. As a result of this, engineering problems have become intricate and require interdisciplinary approaches to unveil novel solutions (Frodeman, 2010). Teaching and communicating these problems may even get more complex since such practices require great effort to produce adequate content and materials. Policymakers have been actively addressing “digital education” to encourage a paradigm shift in education with high-quality, inclusive and sustainable tools (European Commission 2020).

Computational fluid dynamics (CFD) simulations are physics-based numerical tools heavily applied in engineering design and analysis to solve problems in a time cost effective fashion. CFD comprises not only hard skills but also soft and cognitive skills such as critical thinking, decision making, creative problem solving and time management, together with advanced spatial reasoning skills (Tian, 2017). Therefore, the educational use of CFD simulations may be essential to prepare young students for real-life settings.

CFD simulation tools can be directly applied in engineering education. However, they bring several challenges in the implementation and operation for both students and lecturers. First, the educational use of CFD simulations may be challenging due to an expert-centric user experience in conventional simulation and post-processing environments. It requires complex skills to perform CFD simulations and interpret obtained results to make justifiable decisions. Second, learning in conventional simulation environments often happens via learning by doing on 2D desktop settings. Such environments do not comprise any assistance except help options relevant to the usability of the tool. Learning by doing is criticized by educational scientists and found less efficient than traditional instructional designs (Kirschner et al., 2006). Third, CFD simulation data can be utilized in the existing educational context in the form of video clips and images. However, in such practices, students cannot directly interact with simulation data, thereby merely experiencing a passive participation. Finally, lecturers should provide sufficient supportive information to help students understand the basics of the problem solved by CFD simulations, as well as design adequate instructions. All these can hinder the utilization of CFD simulations and simulation data in education in an effective and interactive manner.

The assistance of virtual reality in support of immersive learning has been a hot topic in engineering education. Immersive virtual reality learning environments may reduce the entry barrier to cognitively complex learning subjects (Soliman et al., 2021). These environments can positively trigger cognitive skills with advanced spatial interactions and easy-to-access technical content, as well as behavioral aspects of learning such as attracting and motivating students (Coban et al., 2022). This might open gates for user-friendly, high-quality complex learning environments assisted with CFD simulations.

Developing VR environments with physics-based high-fidelity engineering calculations can create value to facilitate high-quality immersive and interactive

educational tools (Kumar et al., 2021). Likewise, visualization of CFD data in VR can enable a better interface than 2D screens to work with complex datasets in the context of scientific visualization (Vézien et al., 2009). Many researchers have investigated the integration of CFD simulation data in immersive technologies, mainly targeting technical challenges (Berger & Cristie, 2015; Li et al., 2017; Solmaz & Van Gerven, 2021; Yan et al., 2020). Given the complexity of CFD workflow and methodology, immersive visualization of CFD simulations in VR itself is not sufficient for non-expert users such as engineering students. The learning environment should be adequately structured considering relevant components of instructional design. It also appeared that more research on evaluations of prototypes including human factors is essential to provide better applications and unlock the potential of these tools for students to policymakers (Li et al., 2017; Su et al., 2020; Takroui et al., 2022; Yan et al., 2020).

Little is known about the educational use of VR with CFD simulations and it is not clear what factors should be taken into account in the design, development and evaluation phases. In this study, we present the VR application “Virtual Garage” which is a holistic immersive virtual reality experience to educate students with a real-life engineering problem solved by CFD simulation data. Our ultimate goal is to obtain evidence that helps to address these research gaps in support of developers and practitioners to design and evaluate immersive learning environments with CFD simulations.

## 1.1 Related work

### 1.1.1 Visualization of CFD data via immersive technologies

Previous studies have evaluated case-specific visualization options with target users (Baheri Islami et al., 2021; Behrendt et al., 2020; Deng et al., 2021; Gan et al., 2022; Huang et al., 2017; Lin et al., 2019; Logg et al., 2020; Marks et al., 2017; Wei et al., 2021). In another research, visualization of fluid flow in immersive environments was evaluated with a quantitative approach to understanding nonexperts’ interaction and experience. Statistical analysis was also performed to test different hypotheses to pinpoint significance (Christmann et al., 2022). Studies collected data and assessed features with a very limited number of participants. In general, interviews are reported to validate and improve visualization features such as colors, visual representations and interaction with simulation data. Conclusions were drawn upon the features to make simulation data more intuitive and accessible for non-experts in AR and VR.

Likewise, a recent study investigated the effect of immersion in the visualization of blood flow in the vascular systems by comparing a 2D screen, a semi-immersive screen and a VR head-mounted display (HMD) (Shi et al., 2020). Both qualitative and quantitative analyses were performed to evaluate users’ performance and intuition. VR HMD was found innovative and supportive to build cognitive abilities to visualize simulation data and manipulate geometry for different simulation settings. Another work carried out a media comparison study between desktop and VR HMD.

A subjective evaluation was made to assess accuracy, experience and graphics with 8 participants. It was indicated that VR can increase the likeliness of experience and graphics (Yan et al., 2020).

A growing body of literature has evaluated the visualization of CFD data in augmented reality (AR) and virtual reality (VR). Neither instructional design nor assisting educational content is utilized in the digital applications. User assessments were fundamentally conducted through qualitative analysis to get more insight into the cognitive outcomes of the immersive visualization of CFD data. Quantitative methods and standardized tests were occasionally considered by researchers.

### 1.1.2 Educational use of CFD data in immersive learning

In recent years there has been limited interest in the development of learning experiences with instructions in immersive learning environments with CFD simulations. A preliminary attempt was to implement a VR application with CFD data in a master course (Wehinger & Flaischlen, 2020). According to the qualitative assessments, even though negative aspects of VR such as simulator sickness and cost were reported, students overall showed a positive attitude towards the immersive experience. However, several key challenges such as how to tackle downsides to make the implementation easier for students remained unsettled. In another attempt, a VR environment with CFD simulations and instructions was developed to educate farmers and decision-makers in a greenhouse (Lee et al., 2022). A preliminary qualitative analysis was carried out, through which several improvements were processed in the user experiences such as supportive and procedural information, and also the design of a tablet-shaped graphical user interface (GUI). The assessment was only superficially reported and not profoundly discussed. Surprisingly, only one study considered the utility of standardized tests in user assessments (Asghar et al., 2019). A usability study with the System Usability Scale (SUS) was carried out on the VR application to educate gas engineers on the potential gas leaking scenarios. However, no perspective was given on instructional design and other aspects relevant to immersive learning environments.

A recent set of studies investigated the learning effect of a VR application to teach the basics of fluid mechanics and CFD (Konrad & Behr, 2020; Boettcher & Behr, 2021). A quantitative methodology is applied with self-reported metrics to measure understanding and learning, and compare it to conventional teaching. The study claimed that CFD simulations in educational VR positively affected learning and increased students' interest in the content. It was also useful to help students comprehend mathematical equations and spatial reasoning to work with complex CFD simulation data. Subjective metrics were mostly utilized, and no statistical analysis was performed to show significance.

Finally, despite not using immersive technologies, we found another study worthwhile reporting since it compares real-time interactive and non-interactive CFD simulations in a user-friendly learning environment for non-experts (Wang et al., 2021). Both performance and user experience are statistically analyzed through self-reported quantitative metrics. The study indicated that interactive simulations were less demanding than non-interactive ones. Interactive simulations also encouraged

users to explore more parameters. Despite this, the overall task load did not change between interactive and non-interactive simulations. In both cases, since participants were novices to fluid dynamics, challenges were observed in the interpretation of CFD results, such as understanding water flow from visually represented CFD data. This affirms the importance of instructional design in learning environments with CFD simulations.

Recently published reviews have provided a broad list of challenges that are dramatically impeding the adaptation of VR in engineering education. A summary of conveyed challenges is available in the Supplementary Information. In summary, many attempts have been made to assess technical and human factors with user studies including usability, simulator sickness, self-reported scales and learning. Research mostly focused on users' interaction with simulation data. However, most of the studies very primitively reported on important metrics and did not adequately process and analyze research data. Studies were also performed with a very limited number of participants, who were not – in general – truly representative of the target audience. Some applications also directly concentrated on learning outcomes instead of evaluating the quality of the digital environment before measuring the learning. Hence, very little is still known about the importance of human factors in educational VR experiences with CFD simulations. No immersive learning experiences with CFD simulations have been encountered in the literature. Neither learning theories nor instructional design models have been reported by developers.

## 1.2 Objectives and significance

Recent evidence revealed that each VR application has a unique user experience with regard to simulator sickness (Palmisano & Constable, 2022). We can arguably relate this to other human factors in VR since every VR application is composed of various custom design elements such as GUI, instructional design and digital content. An analogous approach has been suggested based on Maslow's hierarchy of needs to systematically evaluate VR applications from the ground up (Tehreem et al., 2022). A preliminary evaluation of human factors is key to unlocking the potential of immersive learning environments to detect underlying and hindering effects before focusing on learning outcomes.

In this study, we evaluated the effect of CFD simulations in an immersive learning environment using the Virtual Garage. It is composed of two subsequent modules; Module#1 and Module#2. The former is a procedural learning environment to teach interaction in VR and provide supportive information on content with relatively easy tasks to be completed by users. The latter is an assessment environment mainly comprising data interpretation, problem-solving and decision-making by using pre-computed CFD simulation data. Even though both modules cover several learning content and tasks, it would appear that they can context-wise quite differ. We, therefore, analyzed usability, user experience, task load and simulator sickness by performing a pairwise comparison between the modules. The following research questions (RQ) are purposefully formulated:

- RQ1: How does system usability change between the modules?
- RQ2: Is there any significant difference in task load between the modules?
- RQ3: How does the content affect the user experience between the modules?
- RQ4: Is there any significant difference in simulator sickness between the modules?
- RQ5: Is there qualitative feedback that may help further interpret quantitative results?

## 2 The Virtual Garage concept: What matters?

The Virtual Garage is a task-centered educational application composed of a holistic immersive virtual reality experience to educate students through a real-life engineering problem solved with CFD simulation data. It is the first-ever learning concept blending virtual reality, instructional design and CFD simulations. It stems from the frameworks of Technological Pedagogical Content Knowledge (TPACK) (Harris et al., 2009) and Substitution, Augmentation, Modification and Redefinition (SAMR) (Hamilton et al., 2016). Both frameworks give a theoretical foundation for the design and development of digital learning environments. Only making technology and content available for learners is not the best practice for immersive learning environments. Suitable pedagogical approaches, such as instructional design models, should be implemented to leverage their quality and value, and to eventually come up with sustainable digital solutions in education. The following subsections detail the applied methodologies as providing foundations for the Virtual Garage concept.

### 2.1 Pedagogy: Instructional design matters

Working with CFD simulations can be a complicated procedure requiring both technical skills and competencies while covering the entire cognitive spectrum in the revised version of Bloom's taxonomy of learning (Tian, 2017). The taxonomy is a prominent framework to effectively identify the learning outcomes. It also categorizes and classifies cognitive skills to reach the utmost complexity in the learning process (Anderson et al., 2001; Krathwohl, 2002). Hence, it is extremely demanding for engineering students to adequately perform and learn in simulation-driven learning environments. A learning environment with simulation data should be properly designed not only concerning data visualization but also supportive and procedural information to sufficiently guide learners to prevent cognitive overload. According to the Cognitive Load Theory, instructional design is an essential tool to develop learning experiences that the human brain appropriately manages (Sweller et al., 2019).

A good instructional design should consider both cognitive load theory and multimedia learning principles, and accordingly adapt them to learners (Mayer, 2014). CFD data is inherently multimodal and composed of varying multimedia such as graphs, charts, numerical data, 2D colored contours and 3D volumetric data

with colormaps. Therefore, in order to prevent cognitive overload, it is imperative to apply multimedia learning principles to effectively communicate CFD results with nonexperts. Among several instructional design models, the four-component instructional design (4C/ID) is being increasingly applied in learning environments for complex learning (Sweller et al., 2019). 4C/ID deals with a task-centered complex learning environment accompanied by real-life tasks while incorporating design principles in multimedia environments (Merriënboer & Kirschner, 2018). From this perspective, the 4C/ID model is a good fit to design immersive learning environments with CFD simulations. Therefore, we meticulously adopted advised steps by the 4C/ID in order to conceptualize and design the Virtual Garage.

The 4C/ID model comprises four major components; learning task, supportive information, procedural information, and part-task practice. Learning tasks are whole-task assignments based on real-life problems that are supported by supportive and procedural information. Supportive information is the theory to provide non-routine aspects of the domain and task that learners can unconditionally reach. Procedural information is timely and organized instructions to help learners to perform routine aspects during the learning task. Part-task practices are practice items to assist learners to gain a high level of automaticity. The components are customized with several principles to orient them with specific learning objectives. A broad spectrum of principles with explanations is available in the relevant design guideline (Kester & Merriënboer, 2021). Details on the implementation of the 4C/ID module in the Virtual Garage with the relevant components, principles and examples are available in the Supplementary Information.

Moreover, we also implement a number of gamification elements in the learning environment. A roleplaying scenario is embedded to fortify real-life experience and immersion (Plass et al., 2020). In this case, users are treated as engineers who are collaborating with stakeholders to solve a problem with CFD in an industrial case study. This may help increase immersion from a pedagogical point of view. In addition, playful learning can increase students' engagement and positively affect emotions (Plass et al., 2020). For that reason, we create playful tasks to provide interactive and joyful experiences such as a puzzle game to grab and correctly locate 3D components of a reactor in a puzzle.

Both formative and summative assessments are made throughout the learning experience. As an example of formative assessment, in Module#1 after a certain amount of supportive information students are asked to complete a puzzle by matching keywords to correctly formulate the objectives of the case study. Similar formative assessments are distributed in the modules. Also, we incorporate a summative assessment methodology based on time, errors, achievements and decisions in the entire application. Overall performance is scored based on an assessment rubric with intended learning outcomes. At the end of their experience, students can observe their performance to make self-evaluations using a learning analytics dashboard. They can also retrospectively move into Module#2. After the VR session, we schedule a dialogue with students to let them reflect on their experience, which may help to form a long-term memory as recently proposed by researchers (Klingenberg et al., 2020).

## 2.2 Virtual reality: Immersion and interaction matter

Immersion and interaction are primary features of VR leading to behavioral, affective and cognitive states of users (Petersen et al., 2022). In the Virtual Garage concept, we utilize VR to deliberately give a boost in these states. While the VR experience can be engaging, motivational or fun; it can also help develop spatial understanding and reasoning skills. This is well in line with our objectives in the development of the Virtual Garage concept. The Virtual Garage is purposefully designed for a 6-degree-of-freedom (DoF) VR experience to unlock the full potential of immersion and interaction in the virtual environment.

In terms of user experience, we targeted a holistic VR experience to provide a standalone learning module, such as learning nuggets in microlearning (Horst et al., 2022), longer than traditional ones but similarly focusing on a case-specific learning activity without any external intervention. In particular, a recent framework has classified VR experiences that vary throughout a continuum between atomistic and holistic experiences (Rauschnabel et al., 2022). Researchers drew a distinction between atomic and holistic models in terms of user experiences. They claimed that holistic VR may enable higher hedonic quality, in contrast to higher pragmatic quality for atomistic VR (Rauschnabel et al., 2022). This should be evaluated by applying relevant metrics and scales in user studies.

Interaction and immersion of VR should be cautiously designed to mitigate simulator sickness. In the field of VR, simulator sickness refers to any physical and mental symptoms affecting users' well-being either during or after the VR experience (Caserman et al., 2021). Literature has recently been bombarded with research on simulator sickness. Several guidelines are made available to lead developers to mitigate simulator sickness in custom VR experiences (Caserman et al., 2021; Kourtesis et al., 2019; Saredakis et al., 2020). Several crucial design aspects were taken into account in the Virtual Garage concept as presented in the Supplementary Information.

## 2.3 CFD simulations: Content matters

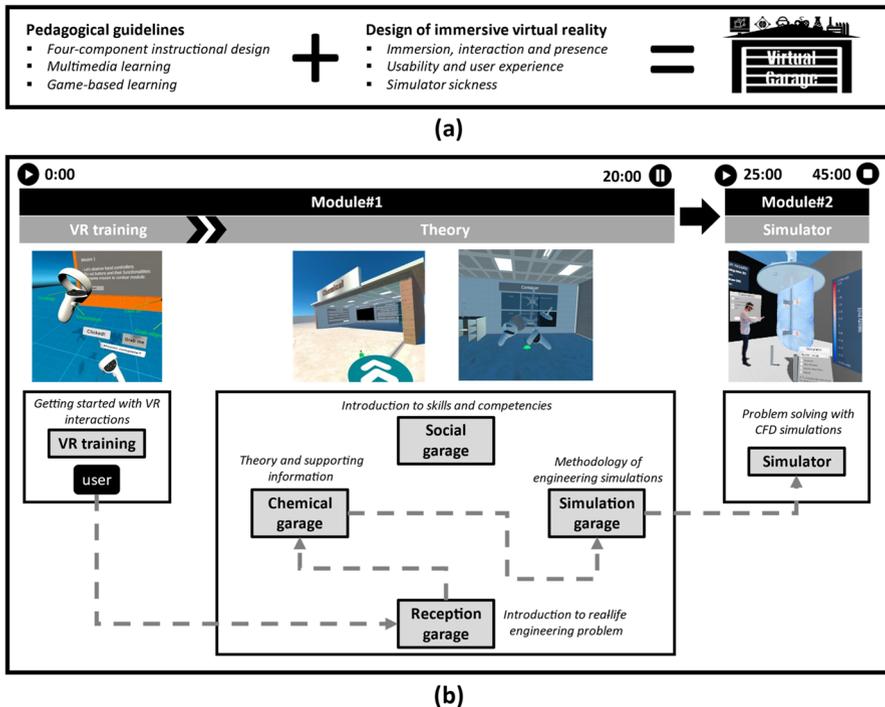
Choosing CFD content that makes sense for students is an important aspect of the immersive learning experience in the Virtual Garage concept. For example, "flow past an immersed object" is a widely utilized case study to train and practice learners in fluid mechanics and CFD. The content is simple, visual and very cheap to calculate. However, the content fails to resonate among learners to teach the methodology and application of CFD since students cannot relate the physical phenomenon to everyday engineering problems. Therefore, from our perspective, it is crucial to find a real-life example and embed simulation content in it, as well as relate the learning experience to the existing engineering curriculum. In the Supplementary Information, we list a series of CFD content to teach intensified processes in chemical engineering based on a content map structured by screening the relevant curriculum.

CFD data should be carefully processed to present meaningful visual and textual data. More accessible colormaps should be considered in visual representations such as batlow (Cramer et al., 2020). Realistic rendering of assisting digital content such as geometry with relevant textures may be helpful. Multimedia learning and relevant principles should be substantiated in the design of the simulation environment.

### 2.4 Prototype: Concept matters

The previous sections detailed design guidelines utilized in the conceptualization of the Virtual Garage. Figure 1 illustrates a schematic of the Virtual Garage concept with applied pedagogical tools. The entire VR experience is divided into two subsequent sections, namely Module#1 and Module#2.

Module#1 is composed of VR training and theory sections. The former aims at teaching hand controllers in VR. The experience resembles a playground where users practice interactions throughout playful tasks. It includes only in-game formative assessments to be completed by users before proceeding with the actual learning content. The ultimate goal of VR training is to get users familiar with hand



**Fig. 1** The Virtual Garage concept including a list of design guidelines utilized in the conceptualization (a); schematic illustration of modules of the Virtual Garage in detail. The user’s move is represented with dashed lines (b)

controllers and interactions in VR, thereby negating the effect of any cognitive load due to the use of technology. The latter is the section where the theory is presented with supportive information together with preliminary tasks to evaluate users throughout a formative assessment scheme. Users consume the content and complete playful tasks in the learning process. Users are instructed to move between three garages to get informed on the varying aspects of the learning content. First, users are born in the “reception garage” where they are welcomed and introduced to the terminology, problem and objectives with simple tasks. Following this, they are instructed to move into the “chemical garage” to learn about the theory and the case study including interaction 3D visualization of geometry with accompanying tasks. Finally, they come to the “simulation garage” to learn fundamentals on the methodology of CFD with case-specific instructions. The Module#1 can provide an engaging meaningful activity that methodically resembles a real-life experience. Because the learning experience covers varying aspects of engineering, it is imperative to let users make distinctions between these aspects using the garage concept. Moreover, we originally designated a “social garage” to highlight engineering competencies and transferable skills such as critical thinking, teamworking, decision-making and so forth. We also aimed at providing an environment to interact with users in the virtual space, such as multiplayer in the context of the metaverse. However, in the current state of the prototype, this garage is inactive due to concerns over exposure time.

Module#2 is the assignment with CFD simulations to solve problems introduced at the end of Module#1. We followed a simplistic approach to designing a greyish virtual room – as a simulator – to get users focused on simulation data instead of distracting them with surrounding irrelevant digital content. Unlike Module#1, users can move only in a confined area mostly around simulation data to make logical interpretations. Users work with pre-computed CFD simulation data to interpret findings and make decisions to satisfy stipulated constraints in the learning activity. Having filled the decision-making panel popping up at the end of Module#2, users can compare their answers to the desired ones, and the entire learning experience comes to an end. Optionally, before quitting the Virtual Garage, users can also retrospectively turn back to simulation data to comprehend the reasonings behind the desired answers. A summative assessment weighs the learning outcomes in Module#2. As an extension to the summative assessment, we see the added value of conducting an oral interview either immediately after or at another moment. In this way, lecturers can get more insight into the learning experience by having an internal evaluation of the learning material and activity. In addition, a learning analytics dashboard can be entailed in the learning experience to timely present overall performance. Both sections take approximately 20 min to complete. Users are strongly advised to take a five min break between modules by taking the VR headsets off, thereby alleviating simulator sickness.

Figure 2 demonstrates screenshots from different scenes of the Virtual Garage application. It is composed of 15 conceptual case studies – microlearning learning nuggets – to provide immersive learning experiences on complex engineering topics, relatively focusing on mixing and process intensification in chemical engineering. 2 out of 15 case studies are prototyped and ready to test with

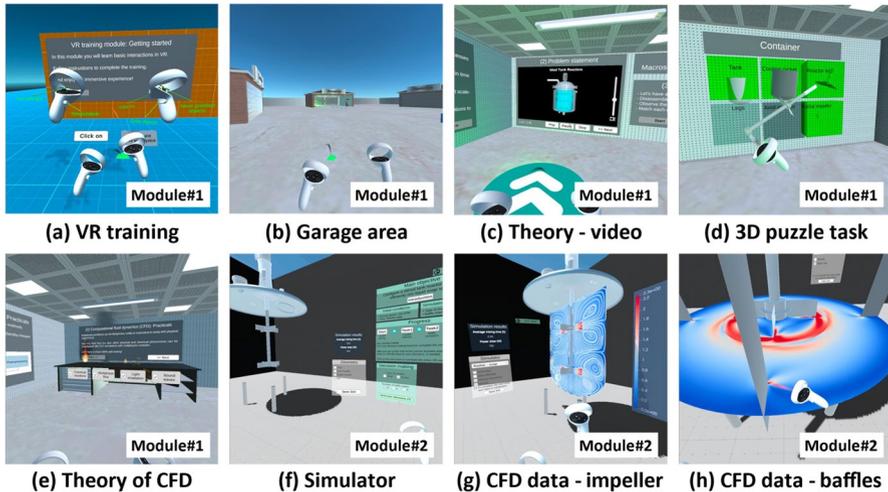


Fig. 2 Screenshots from the Case#6 in the Virtual Garage

students. In this study, we perform user studies with the case study “Case#6: Design a stirred tank reactor”. Users are asked to configure a stirred tank reactor to efficiently mix liquid soap solution using design and operating parameters. Several learning objectives are given to deepen their knowledge of the effect of viscosity, type of impellers, rotational speed and baffle plates, which are detailed in the Supplementary Information. The Virtual Garage is developed with Unity Game Engine and deployed on Meta Quest 2 VR headsets. In the Supplementary Information, a walkthrough video of Case#6 on YouTube and VR software for Meta Quest 2 on GitHub are freely accessible.

### 3 Materials and methods

#### 3.1 Participants

Experiments were conducted at the department of chemical engineering at KU Leuven. Participants filled in a pre-test before the experiment to score their interest and experience in VR, CFD and gaming together with demographic info including gender, age and educational level. As illustrated in Table 1, 24 participants (20 male and 4 female) from the graduate school took part in the testing, of which 17 were Ph.D. students and 7 were Master students. No participants dropped out during the testing. Participants’ age varied from 22 to 31 ( $M = 26.54$ ;  $SD = 2.77$ ). Participants had diverse interests and experiences in VR, CFD and gaming from novices to experts. No one had previously used a virtual reality application with CFD simulations.

**Table 1** Participant demographics ( $n=24$ )

Characteristics	Items	Number	Percentage (%)
Gender	Male	20	83
	Female	4	17
Age	22–24	6	25
	25–27	7	29
	28–31	11	46
Current education	Master	7	29
	PhD	17	71
Any experience in CFD	yes	8	33
Any experience in VR	yes	17	71
Any experience in CFD with VR	yes	0	0
Like gaming	yes	15	63

### 3.2 VR setup

We developed and utilized the Virtual Garage application to evaluate usability and user experience in different modules of the application to understand and improve the overall quality. The Virtual Garage is composed of two subsequent modules, Module#1 and Module#2, as shown in Fig. 1. Module#1 is composed of VR training and theory sections. Module#2 is the assignment with CFD simulations to solve problems introduced in the theory. To develop the Virtual Garage, we utilized the Unity game engine with several built-in and external packages. CFD simulations were calculated in a workstation with OpenFOAM and COM-SOL v5.6, thus facilitating a co-simulation pipeline. CFD simulation data was integrated into Unity using an extract-based data processing approach (Solmaz & Van Gerven, 2021). The VR experience was deployed on Meta Quest 2 VR headsets.

### 3.3 Experimental methods: Procedure and conditions

We pursued a mixed methodology to analyze, interpret and critically reflect on measured scales. Several standardized tests were employed together with self-reported questionnaires and a semi-structured interview. A five-point Likert scale was utilized in this study. Before the actual testing, we performed a pilot study with 5 internal staff showing diverse backgrounds somehow related to the Virtual Garage concept. The pilot study improved the application and helped in the design and validation of the experimental materials and methods.

#### 3.3.1 Evaluation methods

Literature previously showed proven methodologies to examine the quality of immersive virtual reality applications based on usability, user experience, task

load and simulator sickness (Díaz-Oreiro et al., 2019; Reski & Alissandrakis, 2020). These are crucial quantitative scales to facilitate optimal quality for users before diving into learning and task performance. Most obviously, quantitative analysis has its limitations to interpret underlying factors on overall scores. We, therefore, employed a mixed methodology to dig into findings blending quantitative and qualitative results.

**Pre-test** Participants filled out a pre-test to express their experience in and interest in VR, CFD and gaming along with demographic information. This enabled us to examine the effect of independent variables on various other scales measured throughout the experimental campaign.

**Usability** System Usability Scale (SUS) was chosen to evaluate the ease of use of a system as being a massively applied reliable standardized test in the literature (Lewis & Sauro, 2018). The SUS consists of 10 items to quantitatively analyze the usability of digital products and services by means of learnability, efficiency, memorability, errors and satisfaction. Based on the SUS, we can understand the current state of the VR application in terms of usability, and detect the subscales that can be improved.

**Task load** The task load is an imperative metric to deliver digital content in an appropriate way when it comes to educational settings. Despite its original application for aviation, Nasa-Task Load Index (NASA-TLX) has become a well-known multi-dimensional scale to predict task load from cognitive, affective and behavioral aspects in different domains (Hart, 2006). By using NASA TLX, we aim at comprehending changes in task loads that design parameters may cause.

**User experience** A good user experience is key to producing engaging, interactive products. We applied the User Experience Questionnaire (UEQ) which has high scale reliability and validity, as well as being a broadly utilized scale to measure the pragmatic and hedonic quality of user experience (Schrepp et al., 2017). UEQ has both default and short versions. The default version is composed of 6 scales and 26 items, which requires a certain amount of time for dependent paired tests. The short version only comprises 3 scales and 8 items, and can provide a rapid evaluation to quantify user experience for different variants of the same product such as modules in our study.

**Simulator sickness** The immersive VR can induce both physical and cognitive side effects that negatively influence users' well-being. The Simulator Sickness Questionnaire was utilized in this study since it is the most widely utilized scale to examine undesirable side effects arising from immersive VR (Caserman et al., 2021).

**Self-reported questionnaires** During the pilot tests we identified some metrics that cannot be directly answered by standardized scales. Therefore, participants were asked to subjectively report their experience with technology and content using

self-reported metrics. This can enable us further probe and interpret THE outcomes of standardized tests.

**Semi-structured interview** In addition to the standardized and self-reported questionnaire, a semi-structured interview was conducted to interpret the outcomes of questionnaires beyond their limited quantified results. The interview consisted of multiple items to explore likeness, interest, redundancy and quality of CFD visuals, potential improvements, implementation in education and personal review.

### 3.3.2 Data collection

Prior to the study, an ethical approval (G-2021–4281-R2(MAR)) was obtained from the ethics committee at KU Leuven to comply with standards for the processing of personal data in academic contexts. It strictly adheres to the requirements of the General Data Protection Regulation (GDPR) and other applicable laws issued by governing bodies. Barring interviews, all data were collected online on a laptop PC provided to participants during the experiments. Interviews were conducted at the end of the experiments through which the researcher took notes from oral feedback given by each participant.

### 3.3.3 Testing procedure

Experiments were performed in January and February 2022. At that period, the COVID-19 pandemic was still strictly limiting in-person gatherings. Therefore, national and regional safety measures with regard to COVID-19 were applied during the experiments such as proper ventilation of the testing environment, disinfection of equipment and social distancing. Due to safety measures and a limited amount of hardware, participants individually attended the experiments along with the researcher to guide them throughout the entire testing.

Having participants welcomed at the testing place, a short introduction was given about research, ethics, the Virtual Garage application and intended learning objectives. Following this, a consent form was signed by participants, where necessary instructions were provided on their voluntary participation, experimental procedure, data collection, potential risks and discomforts, anonymity and confidentiality. Participation was voluntary; no incentives were given.

Since we did not concentrate on learning in this stage of the research, participants were advised to ask anything content-wise that they struggled with in VR. The experiment started with the pre-test. Then, they were introduced to the VR hardware with relevant personalization settings and started running Module#1 in VR. Right after the completion of Module#1, participants took the VR headsets off and answered standardized and self-reported questionnaires on laptops while having a short break. Later, they turned back in the VR and completed Module#2. The same questionnaires were answered after the completion of Module#2 together with additional self-reported questionnaires on CFD content, overall experience and user behavior toward technology. A semi-structured oral interview was performed at the end of the experimental session. The entire procedure approximately took two hours

per participant including welcoming and instructing participants, arrangement of VR hardware and software, VR testing, questionnaire filling and interview.

### 3.4 Data analysis

Data analysis was carried out separately for quantitative and qualitative data. In the end, quantitative data was mixed with qualitative ones to further explore and interpret analytic findings. Quantitative data, including standardized tests, were analyzed using the relevant guidelines provided by developers. Statistical analysis was performed so as to determine statistical significance by applying descriptive and inferential statistics, as well as (non)parametric analysis. All the analyses were performed using the R programming language for statistical computing. On one hand, mean (M), median (MD), and standard deviation (SD) were calculated to summarize the characteristics of datasets as a part of descriptive analysis. On the other hand, inferential statistics were utilized to test several hypotheses arising from the research questions. Shapiro–Wilk test was run on all data to confirm the central limit theorem and normal distribution. To compare Module#1 and Module#2, we used the sample from the population. Thus, it was advised to apply a dependent test for paired samples. If the data is normally distributed, we used paired t-test, and otherwise Wilcoxon Signed-rank test. Likewise, to analyze correlations among modules, experience and demographics, it was recommended to use Pearson correlation for normally distributed data whereas Spearman correlation for non-parametric analysis. The results presented in this study were considered statistically significant when probability  $p < 0.05$ , marked as bold and \*.

Qualitative data were analyzed given the concept of thematic analysis (Guest et al., 2012). A codebook was developed to structure interview data, through which themes and subthemes were purposefully created. We also quantitatively summarized the qualitative data to point out frequencies of coded themes and subthemes.

## 4 Results and discussion

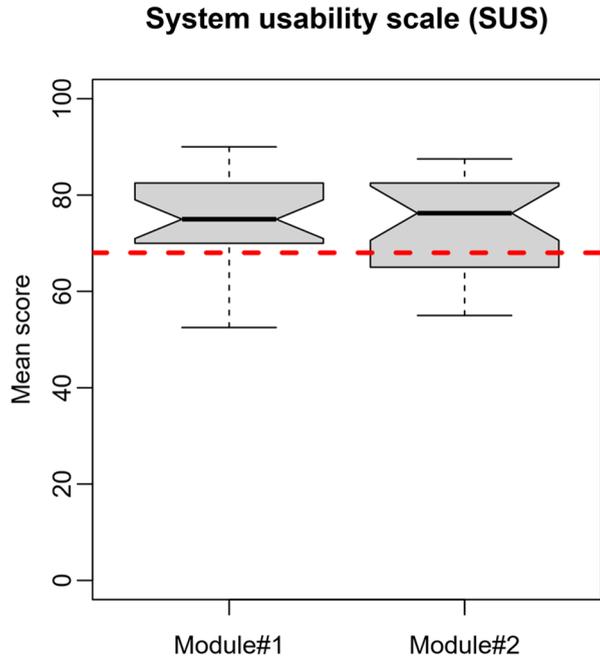
This section concurrently presents and discusses qualitative and quantitative results to facilitate a coherent understanding of our findings.

### 4.1 Usability analysis

Figure 3 shows the mean SUS scores for each module. Both Module#1 and Module#2 were well received by users resulting in mean scores of 74.37 (MD=75, SD=10.01) and 73.85 (MD=76.25, SD=10.16), respectively. A SUS score above 68 is considered acceptable (Lewis & Sauro, 2018) and good (Bangor, 2009).

To examine the differences between modules, a statistical analysis was performed. The Shapiro–Wilk test revealed that Module#1 is normally distributed ( $p=0.252$ ), whereas Module#2 is not ( $p=0.037$ ). Thus, the non-parametric Wilcoxon signed

**Fig. 3** SUS scores. Dashed line thresholds acceptable usability above 68



rank test was chosen. According to the Wilcoxon signed rank test no statistical significance was observed between modules ( $p=0.684$ ).

Moreover, in order to examine the effect of experience and interest on usability, we performed an additional analysis to detect potential correlations. Since the dataset is not normally distributed, we used the Spearman correlation. Only two correlations were statistically significant which were directly related to the SUS score of Module#2. It was revealed that the ones, who find CFD simulation data intuitive on desktop settings, rated higher usability in Module#2 ( $p = 0.037^*$ ,  $r=0.426$ ). Likewise, it was observed that the ones, who find a desktop setting to visualize 3D data

**Table 2** SUS items in modules with relative differences

Item	Module#1	Module#2	Diff (%)
Use frequently	3.67	3.67	0.00
Complex	2.00	1.96	-2.08
Easy	4.08	4.04	-1.02
Need support	2.13	2.29	7.84
Well integrated	4.17	4.17	0.00
Inconsistency	1.58	1.75	10.53
Learn quickly	3.92	4.04	3.19
Cumbersome	2.17	2.25	3.85
Confident to use	3.67	3.75	2.27
Need to learn a lot	1.88	1.88	0.00

difficult, also reported lower usability in Module#2 ( $p = 0.014^*$ ,  $r = -0.493$ ). All in all, these two correlations showed that prior experience with simulation data and 3D models on desktop settings positively reflected users' likeness to Module#2.

SUS score is composed of 10 items as shown in Table 2. Despite the differences among items of modules, neither of them is statistically significant. Interestingly, *need support* and *inconsistency* items were increased by 7.8% and 10.5% in Module#2, respectively. Although these variations are not statistically significant, we further considered blending interview data to possibly clarify what might be the reason behind these increases. The qualitative analysis outlined imperative findings to explain these differences through 3 themes; *CFD scene features*, *interaction with GUI*, and *instructional design*.

In terms of *need support*, some participants—specifically mentioned for Module#2—needed supportive information to interpret and compare simulation data. They proposed help buttons to remind them the important terminologies; for example, *what is the Eddy diffusivity?* Besides, they also demanded a comparison option to put two different data side by side in the same scene, thereby simultaneously comparing differences for two design parameters. Some also mentioned the need for more signaling and procedural information to guide them through multimodal CFD data in Module#2. Eventually, users needed more support in Module#2; therefore, the relevant item in SUS was increased.

The *inconsistency* item was increased in Module#2. Similar to the prior item, some participants mentioned specific features in Module#2 that caused an increase in inconsistency. Notably, users had trouble interacting with the GUI in Module#2, as well as the operations of buttons. In order to switch between different simulation settings, we used the word “initialization” in the dropdown menu, which was not clear for some users despite relevant instructions being provided in the digital environment. It was also revealed that some users were uneasy with the dropdown menu itself, thus finding it unintuitive. Furthermore, the GUI in Module#2 is grabbable, which means that users can grab and relocate GUI in the virtual environment to personalize their experience. There was a button attached to the GUI to manually “save” the latest location. Nevertheless, users forgot to click on the save button, or simply did not understand its function, even though necessary instructions were provided on the same GUI. All in all, some components of the GUI were found unintuitive in Module#2, and this was negatively reflected in the *inconsistency* item of SUS. We believe that using a tablet-like GUI attached to the hand controller can tackle problems with interactions. A similar solution was previously advised in another study without any user assessment of the usability barring some qualitative feedback (Lee et al., 2022).

Thanks to the mixed methodology, it is obvious that the SUS score can be increased by providing more help, supportive and procedural information on the simulation data, as well as making GUI more intuitive and accessible. Not only *need support* and *inconsistency*, but also other items of SUS can be positively affected by these improvements if the same is applied to the entire VR experience.

To sum up, our initial guess on the increases of *need support* and *inconsistency* was simply related to the multimodal simulation data; such as 3D volumetric data, 2D visuals, graphs, charts and analytic data. In contrast to this, our findings showed

that usability was the main reason behind the negative trends in the SUS score of Module#2. It is also worth mentioning that all participants rated the CFD data in Module#2 as intuitive and interactive.

## 4.2 Perceived task load

NASA Task Load Index (NASA-TLX) is applied to quantify the perceived task load of digital systems operated by humans. It measures *mental demand*, *physical demand*, *temporal demand*, *performance demand*, *effort* and *frustration*. We compared the task loads of modules in the Virtual Garage, particularly focusing on how the task load would be impacted in Module#2 due to the use of simulation content and relevant given tasks such as data exploration and decision-making. In this study, unweighted NASA-TLX was used which can provide a sensitive measurement of task load as the weighted counterpart (Hart, 2006). The NASA-TLX scores overall task load in six consecutive categories; low (0–9), medium (10–29), somewhat high (30–49), high (50–79) and very high (80–100) (Hart & Staveland, 1988).

Figure 4 shows overall scores measured with the NASA-TLX test to assess the task load in the modules. Module#1 overall scored medium task load (27.8%) whereas Module#2 scored somewhat high task load (36.8%). Results showed that participants overall found Module#2 more demanding than Module#1, with a 33.4% increase in the average score. Only *physical demand* was lower in Module#2. This trend can be explained by the simplistic design of the simulation environment in Module#2, in which users are centered around simulation data in a confined virtual space. Users were not supposed to move farther distances in the virtual space since the mere focus was to closely explore simulation data to make decisions. In contrast, in Module#1, users move inside the virtual space among garages to consume supportive information and complete tasks to reach Module#2.

Statistical significance was analyzed to deeper examine the comparison between modules. It was observed that the average score was normally distributed ( $p > 0.05$ ) while all of the subscales were not normally distributed ( $p < 0.05$ ). According to the paired sample t-test, the differences between the average score were found to

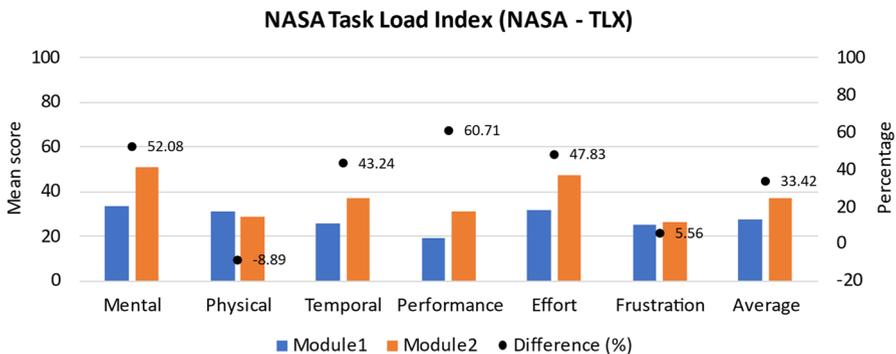


Fig. 4 NASA-TLX test measuring task load in the modules

be statically significant ( $t = -4.786$ ,  $df = 23$ ,  $p < 0.001^*$ ). The task load significantly increased in Module#2, which lends support to our initial guess. Furthermore, the Wilcoxon signed rank test was applied to examine statistical significance in subscales. It was revealed that the *mental demand* ( $p < 0.001^*$ ), *temporal demand* ( $p = 0.025^*$ ), *performance demand* ( $p = 0.023^*$ ) and *effort* ( $p = 0.004^*$ ) were significantly increased in Module#2, whereas the *physical demand* ( $p = 0.388$ ) and *frustration* ( $p = 0.941$ ) were not significantly changed. This clearly substantiates that even though users were more challenged in Module#2, they did not get frustrated at all.

To further interpret the quantitative analysis of the task load, we utilized the interview reports. *Mental demand* significantly increased (52%) in Module#2. Users reported the need for supportive and procedural information. They required guidance to properly interpret the simulation data, and subsequently make logical decisions to solve the problem introduced in Module#2. The tasks in Module#1 were generally simple puzzle games upon the abstract content such as word matching for problem description and geometry dissemble, in which mental demand was far less than in Module#2. Also, a group of participants indicated that they sought to assist features to visualize simulation data such as setting the view direction for different coordinates and angles on click, as well as a comparison mode to put two data side-by-side. It was also requested to make a virtual notebook available for users in Module#2, where they can take notes and return back to them when it comes to decision-making. Otherwise, users either forgot or several times turned back to the same data to memorize their decision. In addition, some users were uneasy with the GUI and its operations due to its unintuitive and inconsistent features. A discussion on this was made in Section 4.1. We believe that all these qualitative reports point to the probability of an increase in *mental demand* in Module#2. On the whole, users needed more instructional support to understand and interpret simulation data along with features to reduce cognitive loads such as a comparison view and virtual notebook in the support of a well-structured decision-making process. Our results are consistent with previous findings in the literature concerning non-experts' interaction with CFD data (Lee et al., 2022).

Similar to *mental demand*, *effort* and *performance demand* were significantly increased in Module#2. We believe that the abovementioned reasons behind the increase in mental demand also played a direct role in increasing *effort* and *performance demand* in Module#2. In addition, the *temporal demand* also increased

**Table 3** Testing correlations on subscales of NASA-TLX; bold asterisk indicates the statistical significance  $p < 0.05$

Correlated items	Module#1		Module#2	
	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>
Mental—Effort	<b>0.010*</b>	0.516	<b>&gt; 0.001*</b>	0.743
Mental—Temporal	-	-	<b>0.039*</b>	0.424
Mental—Frustration	-	-	<b>0.012*</b>	0.502
Physical—Frustration	-	-	<b>0.007*</b>	0.537
Performance—Effort	<b>0.024*</b>	0.460	-	-
Effort—Frustration	<b>0.011*</b>	0.511	<b>0.003*</b>	0.572

in Module#2. Increased task load may result in increased *temporal demand*. It is interesting to note that in Module#2 users had a timer counting down from 20 min, which was not available in Module#1. Interestingly, interview reports unveiled that some participants got stressed due to the timer. This might be one of the reasons behind the increasing *temporal demand* together with *mental*, *effort* and *performance* subscales.

Furthermore, in order to statistically examine the relation among subscales, we performed a test to correlate subscales to each other. Table 3 summarizes statistically significant correlations for modules and relevant subscales with the Spearman correlation for non-parametric analysis. *Mental – Effort* and *Effort – Frustration* are positively correlated in both modules, which provides further evidence on our above interpretations of qualitative data on the increase of *mental demand* and *effort*. In Module#1, *Performance – Effort* is also found to be positively correlated, yet this is not the case in Module#2. In addition, *Mental – Temporal*, *Mental – Frustration* and *Physical – Frustration* are positively correlated in Module#2, but not in Module#1. The results from such analyses should be interpreted with caution. For instance, although the *frustration* was not significantly changed between the modules, it was linked to the *mental demand* in Module#2.

As performed in the usability analysis, we statistically investigated the effect of experience and interest on the perceived task load. Since the dataset is not normally distributed, we used the Spearman correlation. No significant correlation was identified for the average task load score in both modules. Nonetheless, there were significant correlations detected for subscales. The degree of *frustration* in Module#2 is negatively correlated with the ones experienced in 3D modeling on 2D desktop settings ( $p=0.022$ ,  $r=-0.466$ ). The more users are experienced in 3D modeling, the less frustrating Module#2 becomes for them.

Lastly and unexpectedly, there was a significant positive correlation between *mental demand* in Module#2 and expertise in CFD ( $p=0.034$ ,  $r=0.433$ ). In other words, the ones who are experienced in CFD reported higher *mental demand* in Module#2. The interview report might shed light on this unexpected finding. CFD data available in Module#2 were well sufficient to make a final decision to solve the problem. However, the ones experienced in CFD simulations orally reported that they needed more data and data processing options to come to a conclusion such as 3D iso-surfaces, transient results, slices and cut-lines from different coordinates, various scaling options and side-by-side comparison. In general, someone experienced in CFD simulations iteratively carries out data processing work to find optimal post-processing out of the entire CFD dataset. These post-processing are then utilized in the decision-making process. Having in advance final post-processing without any interactive option for reprocessing and data exploration, CFD-experienced users might be challenged due to limited features of the system which is inherently against the workflow they are used to apply, thus scoring higher *mental demand* in the task load. More on the cognitive part of task load could be further factored in a scale specifically developed for immersive VR, thus providing a better understanding of hindering effects (Andersen & Makransky, 2021).

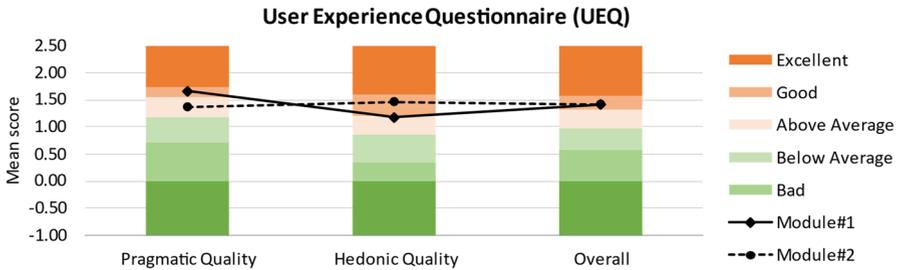


Fig. 5 The UEQ comparing the user experience in modules and benchmark scores

### 4.3 Quality of user experience

Figure 5 shows the UEQ scores of modules together with a benchmark scale to assess the qualities of the user experience. Both modules scored good overall, even though pragmatic and hedonic quality resonated between above average and good. Pragmatic quality covers the items of *supportive*, *easy*, *efficient* and *clear*; while hedonic quality is composed of *exciting*, *interesting*, *inventive* and *leading* items.

The summary of the UEQ scores can be found in Table 4. While Module#1 scored higher in pragmatic quality, Module#2 took the lead in hedonic quality. Statistical analysis with the Wilcoxon signed rank test showed that only the change in pragmatic quality is substantial.

Table 4 Summary of the UEQ scores

Item	Pragmatic Quality	Hedonic Quality	Overall
Module#1	1.66	1.17	1.41
Module#2	1.38	1.47	1.42
Difference (%)	-20.45	20.21	0.55
Statistical significance	<b><math>p = 0.048^*</math></b>	$p = 0.109$	$p = 0.930$

Table 5 Summary of the UEQ item scores; bold asterisk indicates the statistical significance  $p < 0.05$

Items	Module#1	Module#2	Difference (%)	Statistical significance	
Pragmatic quality	Supportive	1.69	1.19	-42.11	$p > 0.05$
	Easy	1.63	1.56	-4.00	$p > 0.05$
	Efficient	1.69	1.69	0.00	$p > 0.05$
	Clear	1.63	1.06	-52.94	$p > 0.05$
Hedonic quality	Exciting	0.94	1.69	44.44	<b><math>p = 0.0197^*</math></b>
	Interesting	1.00	1.44	30.43	$p > 0.05$
	Inventive	1.75	1.75	0.00	$p > 0.05$
	Leading edge	1.00	1.00	0.00	$p > 0.05$

To increase our comprehension of changes in scales, we also took a closer look at items in each scale as summarized in Table 5. No statistical significance was obtained for the items in pragmatic quality, despite the significant change in modules. However, it was observed that items *supportive* and *clear* arguably dropped in Module#2, 42.1% and 52.9%, respectively. Reductions in both may be reflected in the pragmatic quality, thereby significantly reduced in Module#2. As presented and discussed in the previous subsections, some interviewees indicated the need for supportive and procedural information to guide them in Module#2, particularly in the interpretation of simulation data and assistance for decision-making. Similar to this, users also described Module#2 as less clear than Module#1. This may be due to unfamiliar terminology utilized in the GUI and its unintuitive functionalities.

Previously presented and discussed in Section 4.1, the SUS has similar items *need support* and *inconsistency*, which increased in Module#2 by means of usability. This pattern can provide considerable insight into *support* and *clear* items, for which users sought more in Module#2. All in all, the pragmatic quality was significantly reduced in Module#2, which means that users found it less *supportive* and *clear*. It is worthwhile noting that, unlike initially thought, the reduction in pragmatic quality is not because of the simulation data but because of issues in usability and user experience.

Concerning the hedonic quality, also detailed in Table 5, the *exciting* item showed a statistically significant increase in Module#2, whereas no significant difference was observed for other items. The *exciting* item increased by 44.4% in Module#2, alongside the *interesting* item which increased by 30.4%. Yet, due care must be exercised in the discussion because no statistical significance was detected in the overall hedonic quality scale. Several participants expressed their positive feelings on immersive interaction with 3D reactor geometry and volumetric CFD simulation data in Module#2. Some participants even extended their comment to Module#1, for which they hardly ever found the immersive features of VR well exploited as initially expected. However, they were eventually satisfied with Module#2 due to the direct immersive and interactive experience with reactor geometry and simulation data. We presume that these findings can underline just how exciting Module#2 was compared to Module#1, which is basically a structured environment mostly including multimedia to deliver required supportive information with the playful tasks before proceeding with the CFD simulation in Module#2. In addition, all users directly reported that they found the simulation data intuitive and interesting. This positive feeling was clearly indicated in the *interesting* item.

All in all, it would appear that increasing pragmatic quality for Module#2 and hedonic quality for Module#1 can most likely result in better user experiences.

Moreover, our results share a number of similarities with a recently published framework that identifies a telepresence continuum between atomistic and holistic VR experiences (Rauschnabel et al., 2022). The former, in which pragmatic quality is higher, requires less telepresence and is fundamentally aimed at completing a procedural task successfully. This resembles Module#1 in the Virtual Garage. The latter, in which hedonic quality is higher, resembles more a real-life experience giving a high degree of telepresence. Module#2 is not dissimilar to the definition of holistic

experience. Our results showed that users rated higher hedonic quality and lesser pragmatic quality in Module#2 than in Module#1. These findings correlate favorably with the framework (Rauschnabel et al., 2022) and further support the variation in the user experience based on telepresence.

Furthermore, another test was performed to correlate the user experience to demographic data. Spearman correlation was chosen for the non-normal distributed datasets. The pragmatic quality in Module#2 showed a significantly negative correlation with the ones who ranked that using a desktop setting to visualize 3D data is difficult ( $p = 0.013^*$ ,  $r = -0.5$ ). In other words, participants who find desktop settings to visualize 3D data difficult are also challenged in Module#2. Again, experience in 3D modeling might be the important factor here since these participants are more familiar with 3D models and simulation content, and also might have developed higher cognitive skills to work with them.

#### 4.4 Simulator sickness and well-being

The simulator sickness questionnaire (SSQ) calculates total simulation sickness with three different subscales; *nausea*, *oculomotor disturbance* and *disorientation*. A total of 11 participants out of 24 reported sicknesses. 7 out of the 11, who reported sicknesses, did not use any VR experiences prior to this study, and the rest had only one-time experience. Table 6 lists SSQ scores for different subscales and total simulator sickness. All participants successfully completed the entire VR experience without having major discomfort.

10 participants mentioned sicknesses in Module#1, and subsequently seven of them continued reporting similar side effects in Module#2. One participant, who didn't mention any sicknesses in Module#1, reported sicknesses in Module#2. Overall, the degree of sickness significantly lowered in Module#2 compared to Module#1. Even though debates on qualitative analysis of simulator sickness are not yet well settled, recently published research showed that virtual reality environments with an SSQ score below 40 are assumed to be safe in terms of simulator sickness (Caserman et al., 2021). This assumption makes both modules qualified in terms of health aspects.

Why did the sickness scores decrease in Module#2? Design guidelines may have a vivid answer for this trend. In Module#1, users move in VR via teleportation among different virtual buildings, such garages, throughout the experience. This approach was abandoned in Module#2 since we want users to focus on simulation data, thereby less movement to experience in the virtual space. Module#2 also has a

**Table 6** Total SSQ means scores with subscales

SSQ	Simulator sickness subscales			Total simulator sickness
	Nausea	Oculomotor disturbance	Disorientation	
Module#1	17.49	18.95	27.26	23.53
Module#2	8.35	10.42	11.60	11.53

simplistic design in which users are centered on simulation data in a confined space without any distractive surrounding digital assets. These may be the reasons behind the reduced level of sickness. In addition, the scores might be lowered due to the experience they gained in Module#1. No statistical analysis was performed due to discontinuous and segregated datasets.

During the interview session, three participants verbally conveyed perceived physical disturbances caused by the headset, thus affecting their well-being during the VR experience. These three participants also reported sicknesses via SSQ. All of them were novices to VR applications. Simulator sickness can also be triggered by not properly worn VR headsets. Hence, it is imperative to give a short demonstration to participants about how to wear and accordingly adjust the VR headset, and to let them a moment to explore the hardware while finding the optimal setting for their physical comfort. Module#1 comprises a VR training section, in which users do engage with hand controllers and interactions in the virtual space. Though not being applied in the Virtual Garage, the training section can also be utilized to demonstrate what kind of adjustments users can do to properly wear the VR headsets, and to find the appropriate settings. This may further help reduce to the potency of simulator sickness caused by physical disturbance.

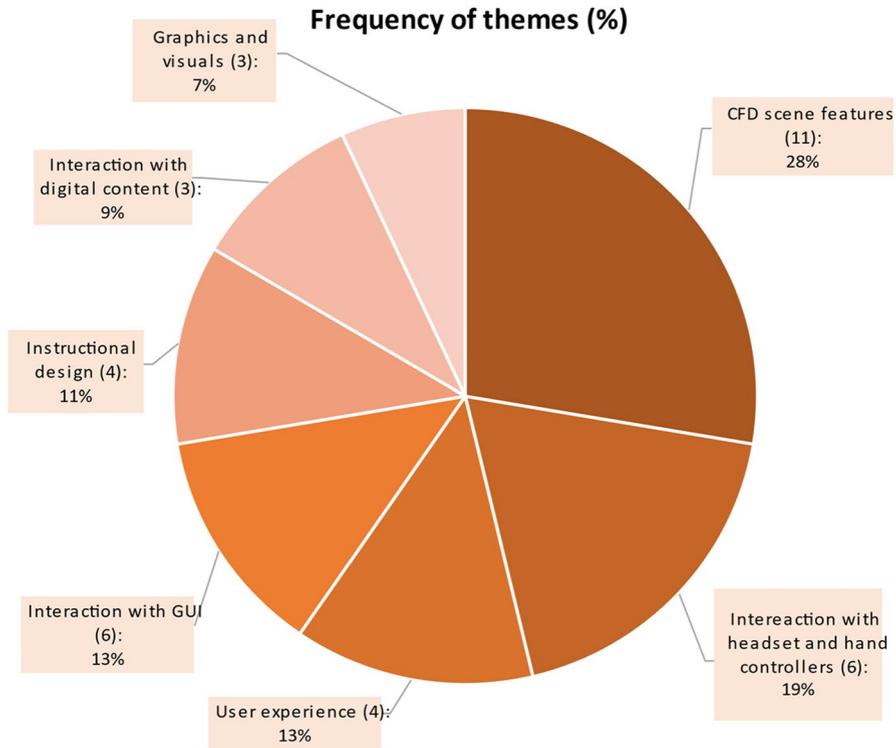
## 4.5 Qualitative feedback

Most of the qualitative data were already blended with quantitative data in previous sections to explore and interpret the reasonings behind user behavior. However, there are still some valuable comments given by participants left that can help us further probe the interesting features of the Virtual Garage. These are discussed in this section.

### 4.5.1 Thematic analysis: Remaining remarks

Seven different themes and 37 subthemes were purposefully generated using a thematic analysis method with a codebook as illustrated by the number of subthemes and frequencies in Fig. 6. Checking on the frequencies of themes in the codebook, some 60% of user feedback is made of “CFD scene features”, “interaction with headset and hand controllers” and “interactions with GUI” themes. This highlights the imperativeness of the qualitative analysis to pinpoint important design parameters. Because immersive virtual reality learning environments are made of truly diverse and subjective elements, standardized questionnaires may not always cover underlying factors.

Interesting remarks can be made on the features of Module#1. The VR training at the beginning of Module#1 was perceived as helpful by the participants. No instructions about operations were given to the participants before the VR experience. Participants learned hand controllers and interactions using the VR training module as a part of a holistic VR experience. This helped to save time and also get participants playfully engaged with VR before dealing with the educational content.



**Fig. 6** Thematic analysis—themes (number of subtheme): frequency (%)

Besides, some participants liked watching content videos in VR and find them engaging despite their negative initial perceptions. Participants also liked and found engaging and intuitive the 3D geometry disassembly puzzle game, as a playful task completed in Module#1 to understand the components of the reactor. Surprisingly, multiphysics animations in the simulation garage were described as being redundant by participants since they were not directly related to the learning content. Our intention to add these animations was to show the capabilities of multiphysics CFD simulation through visual but abstract 3D animations such as smoke propagation and sound waves.

Interviewees also commented on the entire Virtual Garage experience. They found it credible, holistic, fluid and immersive. These comments are in line with the perceived impressions that we want to trigger within the Virtual Garage. Another remarkable comment was on timing and its procedure, which was rated as optimal. Users were given 20 min per module with a five min break in between to diminish simulator sickness.

Furthermore, hand controllers in the Virtual Garage are utilized to interact with GUI and digital content, for example grabbing reactor components. The rest of the operations were conveyed through GUIs available in modules.

Participants found this approach relatively easier than other VR experiences they were previously exposed to. Therefore, they rated interactions in the Virtual Garage simple and easy to control. Although, some users forgot to use the thumb stick, which had the functionality to move a grabbed object in the virtual space, even if it was also a part of the VR training at the very beginning. This implies that at least some hand controller-related procedural information could be conveyed in the modules to remind users of the functionalities of buttons.

Another intriguing finding was about audio instructions in the Virtual Garage. Inherently, we design all audio instructions compulsory to listen to even if a text script was made available in the same scene. Some users found this setting redundant and expressed their intention to skip audio instructions—or at least make them optional to listen—if it is the same as the written script. In addition, some users wanted to change the audio volume and found the default volume set to either quiet or loud. Meta Quest 2 has a button at the bottom of the headset to adjust the audio volume. Neither before the VR experiment nor in the VR training in Module#1 this feature was highlighted to participants. Therefore, our advice here is to give instructions about this button or alternatively provide a headphone which can also help them be isolated from external sounds. Lastly, no negative feedback was collected on background music, for which we were initially concerned that some users might get interrupted.

One participant asked for an option to operate the VR experience while seating. In this study, all participants run the test while standing to mitigate any difference that may arise from this setting. In practice, seating should work since users move in the virtual space via teleportation. A recent study compared both conditions and found no statistical differences between standing up and sitting (Tehreem et al., 2022). In the future edition, users may be accordingly instructed to choose the preferred physical setting. The most obvious advantage of the seating position is that less physical space is required for the VR testing because users would be safe and static in a confined area. Likewise, some participants demanded more precise moving in the virtual space. In this version of the Virtual Garage, users are let freely teleport. However, some stationary points could be spread through the virtual place where users are supposed to be engaging with the virtual content. Instead of moving freely, users click on the stationary point and are directly teleported to the relevant location.

Interesting comments were also revealed on the quality of graphics and colors. Some users expected more quality digital content. Even though Meta Quest 2 is a decent VR device in terms of computing power, it would be quite challenging to increase quality for the sake of realism. It is more important to optimize graphics to enable a fluid VR experience without having any latency and drop in the frame rate. Eventually, it is a trade-off to be optimized between graphics quality and frame rate.

Finally, some users reported on colors, for which they preferred more distinctive colors – or textures—for 3D models and highlighted virtual content. In the Virtual Garage, we carefully chose accessible colors and colormaps for easy interpretation, and people with color and sight disabilities. The same applied to the visual representation of CFD data choosing accessible colormap following a recently published critique on scientific data visualization (Cramer et al., 2020).

### 4.5.2 Relation of experience and interest

VR novice participants found the VR training helpful to learn controllers and interactions in the Virtual Garage. The headset-related physical disturbance was only reported by novices. They also needed more instructions and signaling to navigate themselves through the virtual space.

The ones, who like playing games, described the VR experiences as holistic and fluid. They expected more specific instructions about tasks to be completed, for example, the puzzle game. They demanded more interactive features such as scaling, side-by-side view, and different view options to analyze CFD data. They also preferred the audio instructions to be optional if the written version is made available in the same scene. They liked the personalization of the virtual environment by moving the GUI.

CFD-experienced participants reported that the immersion and interaction are spatial and 3D, and can help develop cognitive abilities to understand and interpret 3D data. They found the playful animations (fire, smoke, etc.) in the simulation garage redundant and cumbersome. Due to their hands-on experience, they also wanted to be exposed to more data such as 3D data, iso-surfaces, transient data, and representative animations such as how the impeller turns – making it more operable, interactive and playful.

### 4.5.3 Integration in the education: Users' perspective

Participants were also asked to give some insights on the implementation of such tools in current educational practices. In general, they agreed on complementary and/or supplementary integration of similar tools to an exercise session for such courses including transport phenomena, microfluidics and broadly engineering design and analysis. Some participants also foresaw its help in teaching CFD and fundamentals of fluid mechanics such as the continuum hypothesis, integral relations for a control volume and dimensionless numbers. One participant detailed a scenario to integrate VR in the current educational settings. The participant proposed integrating it into a group activity preferably during an exercise session. Having an imaginary classroom with 40 students, the participant divided students into groups, forming eight groups of five students. Each group has one VR headset and swaps with each other. In addition, another participant reported that it could be useful for the sake of edutainment to attract young students.

One participant mentioned the potency of the remote utility of the application at home as delivered below:

“It would be useful to download and use it at home. Instead of going to the real plant you use the VR to get there and learn about engineering.”

Using VR only to visualize CFD data was also found an effective interaction to make CFD data intuitive, even if there is no accompanying content in VR. Below are the responses from students highlighting added-value of VR to visualize CFD simulations:

“Very cool for example chemical design problems, do a simulation with COM-SOL, transfer to VR and see the results. This might be helpful there.”

“CFD in VR data helps you understand fluid flow even if you know nothing about the content and simulator.”

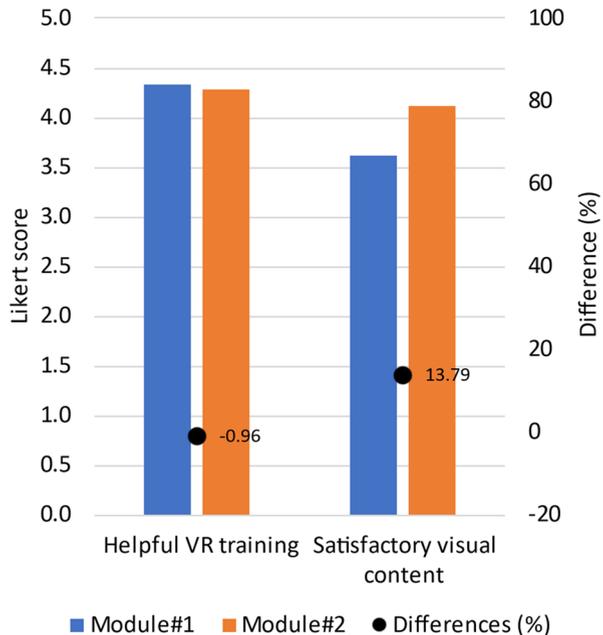
Overall, participants anticipated the added-value of the Virtual Garage and VR technology in their current educational practices. They remained positive towards the implementation; however, they also conveyed doubtful and contradictive comments on their peers’ acceptance of this technology in education.

#### 4.6 Self-reported questionnaire

We further added two self-reported items that may not be directly evaluated in qualitative and quantitative analyses according to the pilot study carried out before the experiments. We asked participants to compare modules by means of the help of VR training at the beginning of Module#1 and satisfaction with visual content, as can be seen in Fig. 7.

A five-point Likert scale was utilized as in the other questionnaires. VR training was found helpful in both modules and didn’t show a significant change between them. In contrast, the likelihood of satisfactory visual content was significantly increased in Module#2 ( $p = 0.0197^*$ ). This substantiates our findings via standardized tests that users rated Module#2 more exciting and interesting, the reason behind the increase in hedonic motivation in UEQ.

Fig. 7 Self-reported items



## 5 Limitations and perspectives

### 5.1 Study limitations

Our study obviously has some limitations. Given that the focus of this study was on the evaluation of usability, user experience, task load and simulator sickness; we couldn't provide any remarks on the task performance and learning effect. This may discourage some readers who seek evidence to justify the utility of VR in engineering education. However, it should be noted that this study is the first step toward assessing a holistic immersive virtual reality learning environment with CFD simulations. Hence, we purposefully concentrated on human factors instead of directly moving into task performance and learning assessment.

Due to the holistic but diverse structure of the Virtual Garage, we couldn't come up with any existing educational practices for comparison. Instead, we compared two modules with and without CFD simulation data to understand its effects on the usability and other measured scales in this study. Alternatively, we could have compared different design features of CFD data in Module#2 such as coloring, scaling option, varying visual representations and instructional design. Several of these features were properly implemented in the Virtual Garage using available literature and guidelines. Nevertheless, our study brought several other features under the spotlight that can help learners to interact with simulation data in VR, for example, side-by-side view selection and a virtual notebook. Furthermore, the assessment part in the Virtual Garage (learning analytics dashboard, retrospective feedback, and post-dialogue) was not yet implemented. Hence, we did not provide any results and discussion on performance and learning assessment. We plan to process a set of those in the next version.

Based on a very small-scale pilot study prior to the testing, we found that a qualitative methodology can help interpret quantitative data. Thus, a concurrent mixed methodology was followed in this study, collecting quantitative and qualitative data. Data collection was made at the same time due to our concerns over reaching out to the same population again. This left some ambiguities and unjustified findings because a larger population brought a diverse dataset to tackle than the pilot study. If the participants are reachable without paying too much effort after the study, it would suit best to subsequently carry out the qualitative data collection based on quantitative results.

We were able to recruit 24 participants for this set of the study. Due to COVID-19-related restrictions, we had to run experiments one by one. Another challenge was the number of available hardware and sufficient physical space to run experiments. All this obviously limited us to work on a small sample size. Despite this, the study showed interesting results to satisfy posed research questions with reliable data analyses. Participants came from diverse backgrounds in terms of experience and educational level. This also helped us properly evaluate the learning environment and find features to be further improved.

## 5.2 Our perspective

The findings of the present study should be critically interpreted and adopted in the design and development of immersive VR learning environments. Given the scope of our objectives, we particularly focused on CFD simulations in immersive VR learning environments. Nonetheless, the design guidelines that can be extracted from our findings are not merely limited to CFD simulations. Educational practices—where 3D modeling, visualization and engineering simulations are being concerned—can substantiate the findings of our research on the design and development of immersive VR learning environments. This can play a crucial role to shorten the design and prototyping time, as well as shedding light on the evaluation of custom design aspects.

Future work is to first consolidate the outcomes of this study and process relevant improvements in the digital environment. On a broader level, in this study, we did not directly measure affective and behavioral factors. However, there is a need for investigation of students' technology use and acceptance of newly adapted digital educational tools. As a part of future works, we are planning to set up another set of experiments to measure task performance, learning and technology acceptance.

There are also a set of twists that remained unanswered and worth investigating. Our findings suggest the following directions for future research. Firstly, researchers may focus on different attributes of CFD data in VR to help nonexperts effectively interpret simulation data. Secondly, more structured and user-friendly authoring tools may encourage practitioners to try and adapt to such digital environments. Thirdly, measuring lecturers' intention to use VR in engineering education seems to be vital to unleashing hindering factors. Besides, comparing in-class and remote use of VR and its effect on learning would be a valuable contribution since remote learning has been increasingly becoming a popular realm. Finally, collaborative features may further be an enabler for remote and social learning given the increasing popularity of metaverse-like educational environments. Not only education but also engineering can also benefit from multiplayer option to effectively and collaboratively communicate CFD simulation data in immersive VR. See the respective entry for a detailed description of the educational use of metaverse (Mystakidis, 2022).

## 6 Conclusion

The focal point of our work is to investigate the learning effectiveness together with cognitive, affective and behavioral aspects of immersive learning with CFD simulations. To perform a reliable evaluation of these metrics, we initially focus on the assessment of the design parameters in the Virtual Garage concept. Hence, the purpose of this paper is to assess human factors as a first step toward the evaluation of immersive virtual reality learning environments with CFD simulations. Our preliminary analysis has shown promising results regarding the design, development and evaluation of CFD simulations in immersive learning. Considerable progress has been made with regard to the quality of the immersive learning experience with

CFD simulations assessing usability, user experience, task load and simulator sickness. Concerning the scarcity of relevant literature to date, our findings and remarks might help developers to find proper guidance in their development journeys.

Future work will focus on improvements in the prototype in the first place. This will ultimately be pursued by a follow-up user study to measure knowledge gain, task performance and technology acceptance. Notably, we aim at measuring technology acceptance due to its comprehensive but structured methodology to figure out what lies behind students' intention to use similar applications by utilizing miscellaneous factors evidenced by literature such as habit, motivation, personal innovativeness, content value and so forth.

**Acknowledgements** The research article is based on a conference paper presented at The European Society for Engineering Education (SEFI) 2022 conference.

**Funding** This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement 812716. This publication reflects only the authors' view exempting the community from any liability. Project website: <https://charming-etn.eu/>.

**Data availability** The dataset analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval and consent to participate** An ethical approval (G-2021–4281-R2(MAR)) was obtained from the ethics committee at KU Leuven to comply with standards for the processing of personal data in academic contexts. Participants signed a consent form before the experiments.

**Conflict of interest** The authors declare no conflicts of interest.

## References

- Andersen, M. S., & Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning*, *37*, 183–196. <https://doi.org/10.1111/jcal.12478>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Asghar, I., Egaji, O. A., Dando, L., et al. (2019). A Virtual Reality Based Gas Assessment Application for Training Gas Engineers. In: Proceedings of the 9th International Conference on Information Communication and Management. ACM, Prague Czech Republic, pp 57–61
- BaheriIslami, S., Wesolowski, M., Revell, W., & Chen, X. (2021). Virtual Reality Visualization of CFD Simulated Blood Flow in Cerebral Aneurysms Treated with Flow Diverter Stents. *Applied Sciences*, *11*, 8082. <https://doi.org/10.3390/app11178082>
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, *4*(3), 114–123.
- Behrendt, B., Piotrowski, L., Saalfeld, S., Preim, B., & Saalfeld, P. (2020). The virtual reality flow lens for blood flow exploration. *Eurographics Workshop on Visual Computing for Biology and Medicine*, 37–41.
- Berger, M., & Cristie, V. (2015). CFD Post-processing in Unity3D. *Procedia Computer Science*, *51*, 2913–2922. <https://doi.org/10.1016/j.procs.2015.05.476>
- Boettcher, K., Behr, A. S. (2020). Teaching fluid mechanics in a virtual-reality based environment. *IEEE Global Engineering Education Conference (EDUCON)*, pp. 1563–1567.

- Boettcher, K., & Behr, A. (2021). Using Virtual Reality for Teaching the Derivation of Conservation Laws in Fluid Mechanics. *International Journal of Engineering Pedagogy (iJEP)*, 11, 42. <https://doi.org/10.3991/ijep.v11i4.20155>
- Caserman, P., Garcia-Agundez, A., GámezZerban, A., & Göbel, S. (2021). Cybersickness in current-generation virtual reality head-mounted displays: Systematic review and outlook. *Virtual Reality*, 25, 1153–1170. <https://doi.org/10.1007/s10055-021-00513-6>
- Christmann, O., Fleury, S., Migaud, J., et al. (2022). Visualizing the invisible: User-centered design of a system for the visualization of flows and concentrations of particles in the air. *Information Visualization*, 21, 311–320. <https://doi.org/10.1177/14738716221085965>
- Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review*, 36, 100452. <https://doi.org/10.1016/j.edurev.2022.100452>
- Cramer, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. *Nature Communications*, 11, 5444. <https://doi.org/10.1038/s41467-020-19160-7>
- Deng, S., Wheeler, G., Toussaint, N., et al. (2021). A Virtual Reality System for Improved Image-Based Planning of Complex Cardiac Procedures. *Journal of Imaging*, 7, 151. <https://doi.org/10.3390/jimaging7080151>
- Díaz-Oreiro, I., López, G., Quesada, L., & Guerrero, L. A. (2019). Standardized questionnaires for user experience evaluation: A systematic literature review. *Multidisciplinary Digital Publishing Institute Proceedings* 31, 14. <https://doi.org/10.3390/proceedings2019031014>
- European Commission. (2020). Digital education action plan 2021–2027. [https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan\\_en](https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en). Accessed 16 Mar 2023.
- Frodeman, R. (2010). Corporate Innovation. In: *The Oxford Handbook of Interdisciplinarity*. Oxford University Press.
- Gan, V. J. L., Liu, T., & Li, K. (2022). Integrated BIM and VR for Interactive Aerodynamic Design and Wind Comfort Analysis of Modular Buildings. *Buildings*, 12, 333. <https://doi.org/10.3390/buildings12030333>
- Guest, G., MacQueen, K., & Namey, E. (2012). *Applied Thematic Analysis*. SAGE Publications Inc.
- Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) Model: A Critical Review and Suggestions for its Use. *TechTrends*, 60, 433–441. <https://doi.org/10.1007/s11528-016-0091-y>
- Harris, J., Mishra, P., & Koehler, M. (2009). Teachers' Technological Pedagogical Content Knowledge and Learning Activity Types: Curriculum-based Technology Integration Reframed. *Journal of Research on Technology in Education*, 41, 24.
- Hart, S. G. (2006). Nasa-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50, 904–908. <https://doi.org/10.1177/154193120605000909>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Horst, R., Gerstmeier, S., Naraghi-Taghi-Off, R., et al. (2022). Virtual reality content creation based on self-contained components in the e-learning domain: Re-using pattern-based vr content in different authoring toolkits. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-022-13362-5>
- Huang, J. M., Ong, S. K., & Nee, A. Y. C. (2017). Visualization and interaction of finite element analysis in augmented reality. *Computer-Aided Design*, 84, 1–14. <https://doi.org/10.1016/j.cad.2016.10.004>
- Kester, L., & van Merriënboer, J. J. G. (2021). Implications of the four component instructional design model for multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *Cambridge Handbook of Multimedia Learning* (3rd ed., pp. 222–262). University Press.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41, 75–86. [https://doi.org/10.1207/s15326985ep4102\\_1](https://doi.org/10.1207/s15326985ep4102_1)
- Klingenberg, S., Jørgensen, M. L. M., Dandanell, G., et al. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology*, 51, 2115–2138. <https://doi.org/10.1111/bjet.13029>
- 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, 342. <https://doi.org/10.3389/fnhum.2019.00342>
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41, 212–218. [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)
- Kumar, V. V., Carberry, D., Beenfeldt, C., et al. (2021). Virtual reality in chemical and biochemical engineering education and training. *Education for Chemical Engineers*, 36, 143–153. <https://doi.org/10.1016/j.ece.2021.05.002>
- Lee, S.-Y., Kim, J.-G., Kim, R.-W., et al. (2022). Development of three-dimensional visualisation technology of aerodynamic environment in fattening pig house using CFD and VR technology. *Computers and Electronics in Agriculture*, 194, 106709. <https://doi.org/10.1016/j.compag.2022.106709>
- Lewis, J. R., & Sauro, J. (2018). Item Benchmarks for the. *System*, 13, 10.
- Li, W., Nee, A., & Ong, S. (2017). A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. *Multimodal Technologies and Interaction*, 1, 17. <https://doi.org/10.3390/mti1030017>
- Lin, J.-R., Cao, J., Zhang, J.-P., et al. (2019). Visualization of indoor thermal environment on mobile devices based on augmented reality and computational fluid dynamics. *Automation in Construction*, 103, 26–40. <https://doi.org/10.1016/j.autcon.2019.02.007>
- Logg, A., Lundholm, C., & Nordaas, M. (2020). Finite element simulation of physical systems in augmented reality. *Advances in Engineering Software*, 149, 102902. <https://doi.org/10.1016/j.advengsoft.2020.102902>
- Marks, S., White, D., & Singh, M. (2017). Getting up your nose: A virtual reality education tool for nasal cavity anatomy. *SIGGRAPH Asia 2017 Symposium on Education on - SA '17* (pp. 1–7). ACM Press.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: A systematic approach to four-component instructional design* (3rd ed.). Routledge.
- Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2, 486–497. <https://doi.org/10.3390/encyclopedia2010031>
- Palmisano, S., & Constable, R. (2022). Reductions in sickness with repeated exposure to HMD-based virtual reality appear to be game-specific. *Virtual Reality*. <https://doi.org/10.1007/s10055-022-00634-6>
- Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers and Education*, 179, 104429. <https://doi.org/10.1016/j.compedu.2021.104429>
- Plass, J. L., Mayer, R. E., & Homer, B. D. (Eds.). (2020). *Handbook of game-based learning*. The MIT Press.
- Rauschnabel, P. A., Felix, R., Hinsch, C., et al. (2022). What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior*, 133, 107289. <https://doi.org/10.1016/j.chb.2022.107289>
- Reski, N., & Alissandrakis, A. (2020). Open data exploration in virtual reality: A comparative study of input technology. *Virtual Reality*, 24, 1–22. <https://doi.org/10.1007/s10055-019-00378-w>
- Saredakis, D., Szapak, A., Bircckhead, B., et al. (2020). Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Frontiers in Human Neuroscience*, 14, 96. <https://doi.org/10.3389/fnhum.2020.00096>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4, 40. <https://doi.org/10.9781/ijimai.2017.445>
- Shi, H., Ames, J., & Randles, A. (2020). Harvis: an interactive virtual reality tool for hemodynamic modification and simulation. *Journal of Computational Science*, 43, 101091. <https://doi.org/10.1016/j.jocs.2020.101091>
- Soliman, M., Pesyridis, A., Dalaymani-Zad, D., et al. (2021). The Application of Virtual Reality in Engineering Education. *Applied Sciences*, 11, 2879. <https://doi.org/10.3390/app11062879>
- Solmaz, S., & Van Gerven, T. (2021). Automated integration of extract-based CFD results with AR/VR in engineering education for practitioners. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-021-10621-9>
- Su, S., Perry, V., Bravo, L., et al. (2020). Virtual and Augmented Reality Applications to Support Data Analysis and Assessment of Science and Engineering. *Computing in Science and Engineering*, 22, 27–39. <https://doi.org/10.1109/MCSE.2020.2971188>

- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*, 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Takrouiri, K., Causton, E., & Simpson, B. (2022). AR Technologies in Engineering Education: Applications, Potential, and Limitations. *Digital*, *2*, 171–190. <https://doi.org/10.3390/digital2020011>
- Tehreem, Y., Fracaro, S. G., Gallagher, T., et al. (2022). May i remain seated: A pilot study on the impact of reducing room-scale trainings to seated conditions for long procedural virtual reality trainings. *8th international conference on virtual reality (ICVR)*, pp. 62–71. <https://doi.org/10.1109/ICVR55215.2022.9848222>
- Tian, Z. F. (2017). Teaching and enhancement of critical thinking skills for undergraduate students in a computational fluid dynamics course. *International Journal of Mechanical Engineering Education*, *45*, 76–88. <https://doi.org/10.1177/0306419016674133>
- Vézien, J. M., Ménélas, B., Nelson, J., et al. (2009). Multisensory VR exploration for computer fluid dynamics in the CoRSAIRE project. *Virtual Reality*, *13*, 257–271. <https://doi.org/10.1007/s10055-009-0134-1>
- Wang, M., Férey, N., Magoulès, F., & Bourdot, P. (2021). Interactive simulation for easy decision-making in fluid dynamics. *Eurographics*, *2021*, 4. <https://doi.org/10.2312/egs.20211022>
- Wehinger, G. D., & Flaischlen, S. (2020). Studying computational fluid dynamics in a new dimension with virtual reality. *Computer-Aided Chemical Engineering*, pp. 2041–2046. <https://doi.org/10.1016/B978-0-12-823377-1.50341-4>
- Wei, Y., Orlosky, J., & Mashita, T. (2021). Visualization and Manipulation of Air Conditioner Flow via Touch Screen. *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 430–431. <https://doi.org/10.1109/VRW52623.2021.00097>
- Yan, J., Kensek, K., Konis, K., & Noble, D. (2020). CFD Visualization in a Virtual Reality Environment Using Building Information Modeling Tools. *Buildings*, *10*, 229. <https://doi.org/10.3390/buildings10120229>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary information** The online version contains supplementary material available at [https://github.com/sersolmaz/The-Virtual-Garage/blob/main/Supplementary\\_Information\\_CASE6.pdf](https://github.com/sersolmaz/The-Virtual-Garage/blob/main/Supplementary_Information_CASE6.pdf).

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.