

Fanika LUKAČEVIĆ, Stanko ŠKEC, Peter TÖRLIND, Mario ŠTORGA

Identifying subassemblies and understanding their functions during a design review in immersive and non-immersive virtual environments

© Higher Education Press 2020

Abstract Design review (DR) is a product development (PD) activity used to inspect the technical characteristics of a design solution. Immersive virtual reality (IVR) technology enables the presentation of spatial information and interaction with 3D CAD models inside an immersive virtual environment (IVE). Such capabilities have shown the potential to mitigate the cognitive load needed for the visual perception of spatial information and, consequently, enhance design understanding and DR performance. Thus, an increasing number of studies have explored the effect of IVR technology on DR activities in different domains. However, determining when the implementation of IVR technology rather than a conventional user interface for DRs in mechanical engineering PD projects will be beneficial remains unclear. Hence, a conceptual DR experimental study was conducted to investigate the differences in the ability of engineering students to identify

mechanisms and understand their functions when a design solution for a technical system is presented in an IVE by IVR technology and in a non-immersive virtual environment (nIVE) by a conventional user interface (monitor display, keyboard, and mouse). Data were collected by performing DR tasks and having participants complete a prior experience questionnaire, presence questionnaire, and mental rotations test. Findings of the study indicate that IVR does not support an enhanced ability of engineering students to identify mechanisms and understand their functions compared with a conventional user interface.

Keywords design review, virtual environment, virtual reality, mechanism, function

Received July 24, 2019; accepted February 18, 2020

Fanika LUKAČEVIĆ
Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Ivana Lučića 5, 10 000 Zagreb, Croatia

Stanko ŠKEC (✉)
Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Ivana Lučića 5, 10 000 Zagreb, Croatia; DTU Management, Technical University of Denmark, Kongens Lyngby, Denmark
E-mail: stanko.skec@fsb.hr

Peter TÖRLIND
Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

Mario ŠTORGA
Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Ivana Lučića 5, 10 000 Zagreb, Croatia; Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

This paper reports work funded by the Croatian Science Foundation project IP-2018-01-7269: Team Adaptability for Innovation-Oriented Product Development (<http://www.taide.org>).

1 Introduction

Design review (DR) is a product development (PD) activity used to inspect the technical characteristics of a design (Dieter and Schmidt, 2012) and development costs (Pahl et al., 2007). It is usually held in the form of a meeting (Watanuki, 2010). Specialists from different disciplines are gathered to discuss a design solution and exchange project information (Pahl et al., 2007). According to Dieter and Schmidt (2012), at least three DRs (conceptual, interim, and final) should be conducted during PD because the development time and costs are usually reduced if a design is revised earlier and more often than necessary (de Casenave and Lugo, 2017). Efforts to implement changes in the conceptual phase require significantly lower resources than later ones, and this condition results in the conceptual review having a high impact on a design solution.

The purposes for a DR can be numerous, such as identifying design problems (for example, part, assembly, or performance errors (Watanuki, 2010; Freeman et al., 2016)), clarifying design assumptions, monitoring the state

of a project, and making decisions about the next steps (Mengoni et al., 2009).

Regardless of the specific purpose, a product design specification (also referred to as a technical specification) is commonly used as the basic reference document against which a design solution is compared and evaluated during DR (Pahl et al., 2007; Dieter and Schmidt, 2012). It is formulated at the problem definition phase (Dieter and Schmidt, 2012). However, the product design specification changes throughout PD as an integral part of the transition from one phase to another, in which it is used as an output of one phase and an input to the following one (Pahl et al., 2007). It is often used in the form of checklists of items that should be considered during a DR (such as functions, embodiment, ergonomics, maintenance, and recycling) and are based on requirements (Pahl et al., 2007).

Design understanding is the basis of DR activities (Chandrasegaran et al., 2013; Paes et al., 2017). It can be defined as the ability to comprehend a design solution, acquire knowledge about the design, and make design decisions accordingly (Banerjee et al., 2002). Design understanding generally depends on two factors, namely, the presentation of a design solution (especially design factors, such as geometric shape or functionality, that are specific to applications) and the prior knowledge and experience of design reviewers (Banerjee et al., 2002).

Previous research (Hannah et al., 2012; Chandrasegaran et al., 2013; de Casenave and Lugo, 2017) has argued that the selection of a design representation type affects the ability of users to review, perceive, and understand a design solution. For example, Hannah et al. (2012) have found a strong correlation between design understanding and representation regarding the type of information that users can extract from a specific representation. Various models, such as 2D engineering drawings, 3D CAD model representations, and mock-ups or physical prototypes, can be used to represent a design solution (de Casenave and Lugo, 2017). The selected representation affects the confidence of users in the extracted information about the design solution (Hannah et al., 2012), which is essential for design decision making as an important sub-activity of a DR (Dieter and Schmidt, 2012). However, noticeable gaps exist in the literature on the selection of the most suitable representation of a design solution for DR activities (such as design understanding, communication, or decision making) (Banerjee et al., 2002; Hannah et al., 2012). Design solutions are often presented as 3D models (Satter and Butler, 2015) using CAD program packages (Chandrasegaran et al., 2013). In a conventional DR setup, interaction with a design representation is enabled through a 2D user interface; the monitor screen is used for visualization, and the computer keyboard and mouse are used for navigation and manipulation (Satter and Butler, 2012). Thus, conventional DR is limited by the 2D nature of the interface (Wann and Mon-Williams, 1996; Wolfartsberger, 2019), specifically the diminishing

representation fidelity (Satter and Butler, 2012) and usage efficiency (Chu et al., 1998). To process the representation of a design presented via the 2D interface and to perceive it as a 3D object, users need an enhanced spatial visualization skill (Hubona et al., 1997; Newcombe and Shipley, 2015; de Casenave and Lugo, 2017). This skill enables them to perceive and understand the presented design solution visually. Design understanding is enabled by the visual perception of the intrinsic (orientation and arrangement of parts, size, rotation, translation, and scaling) and extrinsic (location relative to other objects or a reference frame) spatial characteristics of 3D CAD models and virtual environments (VEs) (Newcombe and Shipley, 2015). Visual perception (Paes et al., 2017) affects the spatial skills necessary for spatial comprehension, namely, intrinsic-static (disembedding), intrinsic-dynamic (spatial visualization and mental rotation), extrinsic-static (spatial perception), and extrinsic-dynamic (perspective taking) (Newcombe and Shipley, 2015). Consequently, the improved support of spatial skills and visual perception enhance the performance of reviewers in conducting tasks and activities for which information extracted from a design representation is needed (Wann and Mon-Williams, 1996; Stanney et al., 1998).

Immersive virtual reality (IVR) has been proposed as a solution that mitigates the cognitive load needed for the visual perception of spatial information contained in a VE (Paes et al., 2017). It enables refocusing the attention to the DR task rather than to visualization tools (Hubona et al., 1997; Liu et al., 2014). As a human–computer interaction technology, IVR allows a user to interact with a 3D CAD model inside an immersive VE (IVE) from the first-person point of view (Cruz-Neira et al., 1993; Slater et al., 1995) using a stereoscopic display (Sutherland, 1968) and multimodal interaction (Witmer and Singer, 1998; Coburn et al., 2017). Such capabilities have often been referred to as the fundamental IVR capabilities needed for experiencing presence in an IVE (Hendrix and Barfield, 1996; George et al., 2014). Presence in VE (telepresence) refers to the subjective experience of being in an environment by which the user is perceptually surrounded (Steuer, 1992). It is based on the interpretation of information gathered through the stimulation of senses (Wann and Mon-Williams, 1996). Such definition offsets the virtual reality (VR) experience from the VR technology and allows a comparison of telepresence in various technologies (Steuer, 1992; Witmer and Singer, 1998).

A new generation of IVR technology, which is more affordable, accessible, and immersive than previous generations, has emerged in recent years (Coburn et al., 2017). Various studies have been conducted to investigate the effect of IVR technologies on design activities in different majors and disciplines, including architectural design (Dunston et al., 2010; Paes et al., 2017; Rigutti et al., 2018), construction (Whyte et al., 2002; Bassanino et al., 2010), engineering design (Germani et al., 2009;

Vora et al., 2001; Faas et al., 2014; Wolfartsberger, 2019), and industrial design (ID) (Mengoni et al., 2009). Across various design disciplines, DR has been one of the most frequently investigated activities. Nevertheless, the focus of these studies has mainly been the development and usability of IVR hardware and software (Chu et al., 1998; Freeman et al., 2016; Wolfartsberger, 2019) rather than the effect of IVR technology on the understanding of users of the design and human factors, such as visual perception or spatial skills (Paes et al., 2017). The results and conclusions of published studies are variegated, and they possibly depend on the domains, phases, and activities of design projects during which IVR technologies are applied (Banerjee et al., 2002). Therefore, determining when the implementation of IVR technology rather than a conventional user interface for DRs in mechanical engineering PD projects will be beneficial remains unclear.

The effect of IVR on the different aspects of design understanding should be further researched to enable the creation of a definition of an objective approach for the integration of IVR technology to support DR activities in PD projects. The following are examples of the aspects of design understanding: An understanding of functions, spatial understanding of dimensions and dimensional ratios, and the detection of errors and defects. With such motivation, this study seeks to investigate the effect of IVR technology compared with a conventional user interface (monitor screen, keyboard, and computer mouse) on the ability of engineering students to identify mechanisms as subassemblies of the technical system and to understand their functions. This specific aspect is chosen because an understanding of mechanisms and their functions is one of the fundamental purposes of DR (Pahl et al., 2007; Dieter and Schmidt, 2012). Various high-level objectives of DR, such as evaluating a design and making decisions on how a design and a project should proceed, are built on it (Freeman et al., 2016).

On the basis of a literature review and the conducted experiment, this study aims to answer the following research questions:

(1) Does an IVE (mediated using IVR technology) improve the ability of engineering students to identify mechanisms and understand their functions compared with a non-immersive VE (nIVE; mediated using the conventional user interface of a monitor screen, a mouse, and a keyboard)?

(2) What is the relationship between the presence that engineering students experience in a VE and their ability to identify mechanisms and understand their functions?

Hypotheses defined on the basis of the abovementioned research questions and used for further definition and execution of the experiment are as follows:

(1) An IVE improves the ability of engineering students to identify mechanisms and understand their functions.

(2) A great presence experienced in a VE yields an

enhanced ability for engineering students to identify mechanisms and understand their functions.

2 Research background

This section presents an overview of presence measures in VEs and their relationship with DR performance as an initial step toward an improved understanding of the subsequent experiment and analysis (Section 2.1). A description of IVR setups and a detailed explanation of how visual representations are perceived (Section 2.2) follow. Section 2.3 presents an overview of related literature that has investigated how IVR technology supports DR activities.

2.1 Presence measures

The most common way of measuring the presence experienced in a VE is through subject self-report in the form of questionnaires (Lombard and Ditton, 1997). Different metrics have been used for presence evaluation, which consider technological (for instance, realism, sensory breadth and depth, or a number of sensory input devices) and individual (associated with the automatic perceptual processes, conscious processes, or the mindful direction of attention of users) factors (Steuer, 1992; Banerjee et al., 2002). Several of the commonly used questionnaires are the Presence Questionnaire (PQ) (Witmer and Singer, 1998), Independent Television Commission-Sense of Presence Inventory (Lessiter et al., 2001), Igroup Presence Questionnaire (Igroup, 2019), Temple Presence Inventory (Lombard et al., 2009), and Spatial Presence Experience Scale (Hartmann et al., 2015). An intense sense of presence in a VE has been argued to improve cognitive performance, lead to realistic behavior in the VE, and enable the perception of the VE as a real environment (Aust et al., 2011; Heydarian et al., 2015). Nevertheless, the relationship between presence and DR performance has not been clearly defined. For example, the results of the study in which Banerjee et al. (2002) compares PowerPoint to Cave Automatic Virtual Environment (CAVE)-based DRs shows that participants who have attended the PowerPoint presentation score higher on the questionnaire concerned with design understanding than those who did not attend the presentation. However, the correlation between the presence and the design understanding score was insignificant. In another study, Faas et al. (2014) have investigated the effect of presence experienced in a real environment on the quality of a designed solution (using a device for the vertical displacement of a ball). The results show a correlation between high presence and high- and low-quality designed solutions, whereas low presence is correlated with average quality (Faas et al., 2014). In other studies (Hubona et al.,

1997; Witmer and Singer, 1998; Banerjee et al., 2002), the correlation between high experienced presence and high task performance is significant only in some cases. The results are often associated with prior experience of participants, such as experience using various setups of IVR technology.

2.2 IVR setups and perception of visual representations

Selected visualization technology defines the visual representation of virtual models (Paes et al., 2017). A mental representation of a virtual model has to be produced to enable its perception (Paes et al., 2017). The stereoscopic effect of human binocular vision allows the spatial visualization of the representation and perception of the space depth and virtual model as a 3D object (Hubona et al., 1997; Chen et al., 2012). Using an analogy of human vision, IVR technology uses a stereoscopic display that provides different visual information for each eye through the divided display (Sutherland, 1968). Although stereoscopic display is a basis for a spatial interaction with a visually perceived VE (Schubert et al., 2001; Berg and Vance, 2017), multimodal interaction and synchronized motion in real and virtual environments (Hendrix and Barfield, 1996; Hubona et al., 1997) enable full immersion and enhanced interactivity (George et al., 2014). IVR technology setups that have been commonly used in DR studies are CAVE, Head-Mounted Display (HMD), and PowerWall (de Casenave and Lugo, 2017). CAVE systems use projection onto multiple walls (at least three) that surround users to display representations, whereas PowerWall uses a single large wall (de Casenave and Lugo, 2017). An HMD system employs a device worn on the head of the user, with stereoscopic displays on a short distance from the eyes of the user (Sutherland, 1968). The space requirements and costs of HMDs are significantly lower than those of the other setups (Coburn et al., 2017), which have resulted in the HMDs becoming a common solution for DRs in academic and industrial applications.

2.3 DR activities supported by IVR

Recent empirical studies conducted in an academic or industrial environment using IVR for DR activities have mostly been concerned with the interaction within an IVE, the usability of IVR (Satter and Butler, 2015; Freeman et al., 2016; de Casenave and Lugo, 2017), and design understanding in terms of spatial comprehension and error detection (Bassanino et al., 2010; Dunston et al., 2010; Rigutti et al., 2018; Wolfartsberger, 2019).

Interactions during a DR can be supported by IVR in diverse ways, such as by allowing experienced and inexperienced users to manipulate a 3D CAD model inside an IVE or conducting enhanced ergonomic, aesthetic, and dimensional assessments (Germani et al., 2009). Satter and Butler (2015) have reported that

participants in their study need significantly less time to perform navigation and error-detection tasks common in the engineering design domain when using stereoscopic interfaces than when using conventional ones. Freeman et al. (2016) have compared IVEs with implemented basic (translation, rotation, zoom in, and zoom out) and enhanced (such as hide function, exploded view, and parametric modeling) functions in gear-counting tasks. The availability of the latter has resulted in users having improved design understanding and enhanced confidence in their answers regardless of their accuracy. The intuitiveness of IVR technology refers to the interface and its usability in terms of ease of memorizing or learning the functionalities and controls needed to use it (Chu et al., 1998; Germani et al., 2009). For example, de Casenave and Lugo (2017) have not found significant differences in the ability of users to detect errors in virtual models in VE and IVE. However, the intuitiveness of IVR technology enables them to interact naturally with VE and analyze the virtual model successfully despite their lack of experience in using IVR technology.

The results of the research conducted by Maftai and Harty (2016) indicate that the use of IVR technology enhances and adds to an existing understanding of a design (one example is discovering unexpected issues about a model (Whyte et al., 2002)). The technology use also challenges the understanding of the participants regarding the design and triggers new ways of making sense of it, thus affecting design decisions and development (Maftai and Harty, 2016). On the basis of the responses of participants to a questionnaire on the usage of IVR, Liu et al. (2014) have concluded that IVR helps users gain an improved understanding of and enhanced confidence in a design. Paes et al. (2017) have stated that in architectural design, an IVE leads to an enhanced spatial perception and an understanding of a virtual model in accordance with the high accuracy in the answers of users. Schnabel and Kvan (2003) have explored the spatial perception of users in VEs and reported that the results of an IVE review are thorough and leads participants to an enhanced comprehension of dimension ratio and design. Nevertheless, the most accurate results of model volume remodeling are scored while participants are reviewing the model on a monitor screen. The results of the experimental part of the study that Dunston et al. (2010) conducted in the construction domain have revealed that an IVE is efficient for the design error detection in hospital room layouts, such as determining insufficient clearance among elements or over-dimensioned elements. Similarly, Bassanino et al. (2010) have inspected the accessibility of a room for wheelchair users. Rigutti et al. (2018) have conducted research to compare the ability of users to detect design errors when they are active (first-person view) and passive users. They have concluded that a setup in which every user wears an HMD and actively participates in DR is the most suitable solution for DR and error detection purposes.

Vora et al. (2001) have used IVR technology to inspect an aircraft model for errors. The scholars have concluded that IVR technology can be an efficient tool for visual inspection in a realistic environment.

IVR technology enables the presentation of spatial information that has shown potential to mitigate cognitive load and, consequently, enhance design understanding and task performance efficiency (Liu et al., 2014; Berg and Vance, 2017). An increasing number of studies have accordingly explored the effect of IVR technology on DR activities in different domains, as shown in Table 1. Design solutions reviewed in various studies range from solid figures (such as cube assembly) to technical systems of the fourth level of complexity (such as aircraft or building). The classification of technical systems by their level of complexity is defined in accordance with the theory of technical systems (Hubka and Eder, 1988). Technical systems of the third level of complexity have prevailed in the engineering domain. In construction and architectural design, technical systems of the fourth level of complexity are common.

Numerous studies that have examined an understanding of the functions of technical systems can be found in the literature but not as a constituent part of a DR in an IVE or nIVE. For example, Mckenna et al. (2008) are interested in the effects of virtual and physical dissection activities on the ability of students to identify and describe the function and production method of hand-held power drill components. Eckert et al. (2012) have conducted an experiment

on how engineers understand the notions of functions and functional breakdown in the context of design by modification when using a classical function tree for a comparison of their understanding. Studies that investigate DR activities using IVR have mostly focused on a comparison of the different capabilities of IVR setups, time metrics (Freeman et al., 2016), or decision making in IVE. Observations, interviews, and focus groups are often used to examine these capabilities (Berg and Vance, 2017). Nevertheless, studies often lack detailed elaborations on the results of conducted experiments and the underlying variables that influence them. In most cases, they do not explore the relationship between the results of the DR task performance and experienced presence, prior experience, or spatial skills of engineers and the important factors when investigating design understanding and its aspects (Banerjee et al., 2002). Differences in domains, phases of product lifecycle, applications, design solutions, or study participants may moderate the effect IVR technology has on DR and study results.

The literature review conducted by the authors do not reveal studies that compare the ability of engineers or engineering students to identify mechanisms and understand their functions during DR using IVR technology and a conventional user interface in mechanical engineering projects. The DR solution in an IVE is generally assumed to improve the understanding of users regarding how components are spatially arranged and how each component interacts with others (Berg and Vance, 2017).

Table 1 Overview of the literature that has researched how IVR technology supports DR activities

Article	Design domain	Application ^a	Participants ^b	Design solution ^c
Satter and Butler (2015)	engineering	ED, SP	EE, RS	machine unit (IV)
Wolfartsberger (2019)	engineering	ED	EE	power unit (III)
Rigutti et al. (2018)	architectural	ED	RS	corridor (IV)
de Casenave and Lugo (2017)	engineering	ED	RS	cube assembly
Paes et al. (2017)	architectural	SP	EE	entrance hall (IV)
Berg and Vance (2017)	engineering	UIE	EE	vehicle (III)
Freeman et al. (2016)	engineering	DC	S	gearbox (II)
Maftai and Hartly (2016)	construction	DC	EE	hospital room (IV)
Liu et al. (2014)	construction	DC	EE	energy retrofit (IV)
Satter and Butler (2012)	engineering	ED, SP	S, EE	ship (III)
Vora et al. (2001)	engineering	ED	S	aircraft (IV)
Dunston et al. (2010)	architectural	L	O	hospital room (IV)
Bassanino et al. (2010)	construction	L	O	bathroom (IV)
Mengoni et al. (2009)	industrial	DC	EE	shower (III)
Germani et al. (2009)	engineering	DC	EE	shower, washing machine (III)
Schnabel and Kvan (2003)	architectural	SP	S	cube assembly
Whyte et al. (2002)	construction	UIE	EE	building (IV)
Banerjee et al. (2002)	engineering	DC	EE	vehicle (III)

Notes: a. ED (error detection), SP (spatial perception), UIE (perceiving user interface elements), DC (design context), and L (layout); b. S (students), EE (experienced engineers), O (operators), and RS (randomly selected); c. Level of complexity of a technical system (Hubka and Eder, 1988): I (first level), II (second level), III (third level), and IV (fourth level).

However, the authors have not found empirical studies in the mechanical engineering domain that support this assumption. Therefore, this study investigates an outlined lack of clarity on whether IVR technology enhances the ability of engineering students to identify mechanisms and understand their functions during DRs in mechanical engineering design projects.

3 Research methodology

The study is defined as a quantitative and comparative experimental study with regression analysis and descriptive and inferential statistical testing. Its design is based on the principles and methods described by Robinson (2016) and guided by published experimental studies concerned with the different aspects of design understanding in various engineering domains (Freeman et al., 2016; de Casenave and Lugo, 2017; Paes et al., 2017; Faas et al., 2014; Banerjee et al., 2002).

The study is conducted through the following eight steps: (1) development of experimental variables and definition of their measures, (2) selection of a design solution to be reviewed and a definition of the DR tasks and questions, (3) definition of the experimental setup, (4) pilot experiments, (5) definition of a full experimental procedure, (6) running of experiments, (7) analysis of collected data, and (8) discussion of the results. Steps 1–6 are explained in this section. Step 7 is presented in Section 4, and step 8 is presented in Section 5.

3.1 Experimental variables and their measures

The development of the experimental variables and their measures was guided by the literature on DR and IVR technology along with empirical studies on how visual representations affect DR activities and design understanding aspects (Section 2).

On the basis of the focus of this study (Section 1), the ability to identify the mechanisms of a design solution and understand their functions was a specific application (an aspect of design understanding) and a dependent variable in this experiment. Factors that might affect the above-mentioned ability were independent variables, namely, prior experience, experienced presence in a VE, spatial skill, and the VE where the design solution was presented (IVE or nIVE).

Questionnaires, which represented an effective method for collecting primary data (Robinson, 2016), were used to gather information about the prior experiences and experienced presence of participants in a VE. Questionnaires on prior experiences contained 25 questions divided into three sections, namely, personal and demographic information (such as gender, nationality, and height), prior experience in designing and using CAD tools (concerning IVR technology, 3D CAD software, and DRs), and

contextual experience (for example, the number of folded vehicles and adjustments of seat height). The obtained information was used to analyze the effect of the prior experiences of participants on the results and to support a discussion, as is common in comparative studies (Hannah et al., 2012; Paes et al., 2017).

The experienced presence of participants in a VE was measured by the completion of PQ (Witmer and Singer, 1998); spatial skills were measured through the mental rotations test (MRT) (Peters et al., 1995). The PQ and the MRT are standardized and commonly used measures in related studies (Banerjee et al., 2002; Kalisperis et al., 2002; Vora et al., 2001).

The participants were divided into two groups based on their scores on the MRT. The division was conducted to maintain approximately equal distributions of the MRT scores in both VEs. One group reviewed the design solution only in the IVE and the other one only in the nIVE to follow the preclusion of bias from published studies (Mckenna et al., 2008; Hannah et al., 2012) and avoid the potential influence of a previously conducted review in one of the VEs. The mean MRT score of the group that conducted the review in the IVE was 12 ± 4.5 , whereas that of the group in the nIVE was 14.1 ± 4.7 .

The DR part of the experiment consisted of eight tasks and seven associated questions on the judgment of confidence of participants in their answers. The DR tasks referred to the identification of mechanisms, the understanding of their functions, the spatial understanding of dimensions, and the detection of design errors. The participants were asked to judge their confidence in the correctness of their answers (C_{am}) on an open-ended confidence-rating scale (Stankov et al., 2014) from 0% (absolutely unsure) to 100% (absolutely sure). Thus, confidence judgment reflected the beliefs of the participants in the accuracy of their answers concerning a particular task (Stankov et al., 2014). Within the scope of this study, three tasks and respective questions concerning the identification of mechanisms, the understanding of their functions, and the detection of the mechanisms' functional errors are described in Section 4.

3.2 Selected design solution

A lightweight foldable mobility scooter was selected as a design solution to be reviewed because the alignment with previously conducted studies in an engineering design domain (Table 1) was a technical system of the third level of complexity—a device that consisted of subassemblies (mechanisms) and parts that performed a closed function (Hubka and Eder, 1988). It was developed during the project-based course as a result of a collaboration among mechanical engineering, ID, and electrical engineering students from four European universities and an industrial partner (a UK-based company). Therefore, it was an appropriate proxy for a real industrial project with the

available data needed for a DR, namely, initial product requirements, a developed 3D CAD model, and accompanying technical documentation. The participants were exposed to two modes of the same CAD model in both VEs, which were the scooter in driving mode with main dimensions of 1181 mm × 664 mm × 1054 mm (Fig. 1) and the folded scooter with dimensions of 655 mm × 664 mm × 843 mm (Fig. 2).



Fig. 1 Driving mode model.



Fig. 2 Folded mode model.

3.3 Experimental setup and pilot experiments

The experimental layout is shown in Fig. 3. For a DR in the nIVE, a conventional technology setup was used, namely, a high-performance computer, a 24-inch monitor screen (Monitor screen 1 in Fig. 3), a computer mouse, and a keyboard as interaction devices. The resolution of the monitors was 1920 × 1080 pixels with a refresh rate of 60 Hz. Siemens NX 12® with incorporated tools was used as the software package to review the model in the nIVE.

The participants used the HMD IVR system HTC Vive Pro with the associated 3D controllers to review the scooter in the IVE. The system display consisted of a dual AMOLED 3.5-inch diagonal screen with a resolution of

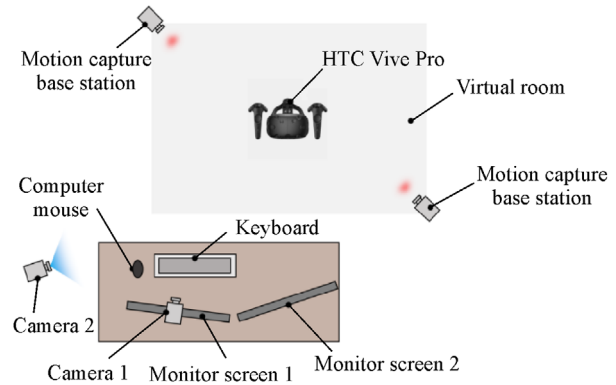


Fig. 3 Experimental layout.

1440 × 1600 pixels, a refresh rate of 90 Hz, and a 110° field of vision. The models were presented in the VRED Professional® 3D CAD package. The IVE was built using VRED OpenVR Scripts to allow the participants to use the model rotation, measuring, and sectioning tools (Fig. 4) in addition to navigation through a 1.9 m × 2 m large virtual room.

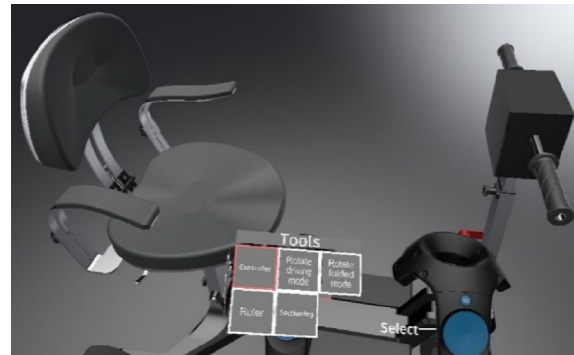


Fig. 4 IVE with associated tools.

Pre-study pilot experiments with five participants were conducted to test the equipment and validate the suitability of the DR tasks and questions. The pilot experiments provided the basis for a definition of the duration of each step of the experiment. From these findings, the full experimental procedure was defined, as explained in Section 3.4.

3.4 Experimental procedure

The experimental duration per participant was approximately 90 min, and it was conducted in accordance with the steps shown in Fig. 5. Participants went through the entire procedure individually. In step 1, the participants were introduced to the experimental procedure and the design solution to be reviewed. In step 2, they were asked to solve the MRT. They reviewed the scooter either in the IVE or nIVE based on the previously described procedure.

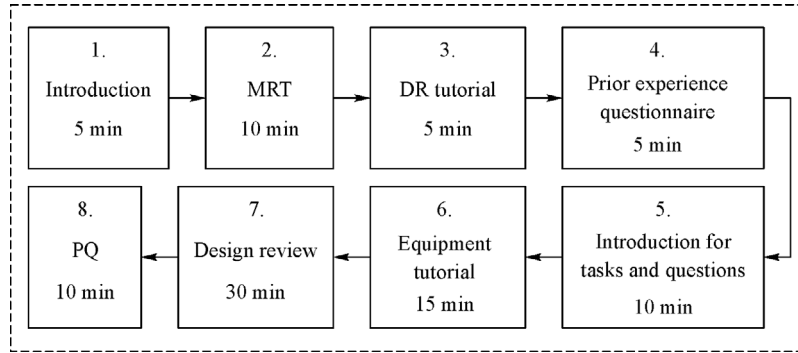


Fig. 5 Experimental procedure.

In step 3, tutorials about the DR were provided. Its main objectives, purpose, and procedure were explained theoretically through two examples. The participants were then asked to complete the questionnaires on their prior experiences. This step was followed by instructions on how to conduct tasks and answer questions. The participants were asked to provide their answers by saying them aloud when performing a DR. In step 6, the participants were introduced to the equipment. First, a video recording of the VE and available tools was shown. Afterward, the participants were asked to conduct four exemplary tasks (they were the same in both VEs) to become familiar with the VEs and tools. The next step was to conduct the DR tasks and answer questions about confidence judgment. The three tasks described within the scope of this work were concerned with (1) the identification and understanding of adjustment mechanisms, (2) the identification and understanding of folding steps, and (3) the detection of functional errors (Section 4). In the last step, the participants were obliged to complete the PQ.

3.5 Running the experiments

The study involved a total of 40 graduate and undergraduate engineering students. Sixteen of them were students from one university, and 24 were from another university. The number of participants was aligned with published studies that had a similar research focus (Ostergaard et al., 2003; Mckenna et al., 2008; Hannah et al., 2012; Vora et al., 2001; Satter and Butler, 2015; Freeman et al., 2016; Paes et al., 2017; Faas et al., 2014), as mentioned in Section 2. The requirement for participating in this study was a course of study in the engineering field. The year of study, specialization, and prior design experience or experience of 3D CAD software and IVR technology usage were not predefined.

4 Analysis of collected data

Descriptive and inferential statistics and regression

analyses were used to analyze the gathered data. R language and environment were used for statistical computing. Descriptive statistics were used to calculate the measures of central tendencies (mean value and median M) and variability (standard deviation and ratio R). Inferential statistic tests were performed to analyze the significance of the difference between the results in IVE and nIVE. For a comparison of two variables, the procedure included (1) a Shapiro–Wilk test, (2) a Levene test, and (3) a t -test or Wilcoxon test (if the Shapiro–Wilk test was not satisfied). Scatter plots were used to depict the results of the DR task performance against the PQ results. Then, a linear regression was used to analyze the correlation among the variables. In the case of categorical variables (number of adjustment mechanisms and folding steps), ANOVA was used to analyze the statistical validity of the model. A Kruskal–Wallis test was conducted for the analysis of the correlation between prior experience and DR task performance.

In total, 33 males and 7 females participated in the experiment. The average age was 22, with a range from 19 to 31. The participants were engineering students in different fields, namely, design and PD (16), ID (15), materials engineering (ME, 7), and sustainable energy engineering (SEE, 2). Their CAD-modeling skills were estimated as good on average (2.4 ± 1.02), where 0 means “none” and 4 means “advanced”. Their CAD-modeling experience was 2.45 ± 1.07 , where 0 means “none” and 4 means “5 or more years”. As presented in Table 2, most of the participants had not conducted a DR prior to the experiment (21) and had no prior experience with an IVR (23).

Table 2 Prior experience

DR experience (time)	0	1–3	4–6	7–9	≥ 10	
Number of participants	21	7	5	2	5	
IVR experience (min)	0	<5	5–15	15–60	150–210	780
Number of participants	23	3	6	4	3	1

4.1 Adjustment mechanisms

In task 1, the participants were asked to identify five adjustment mechanisms, either for height or an angle of the elements of the scooter, and explain their working principles (rotation or translation). The requested adjustment mechanisms were steering bar height (Figs. 6 and 7), tiller angle, armrest height, seat height, and seat angle. The answer was only acknowledged as correct if the adjustment mechanism was followed by a correct explanation of its working principle.

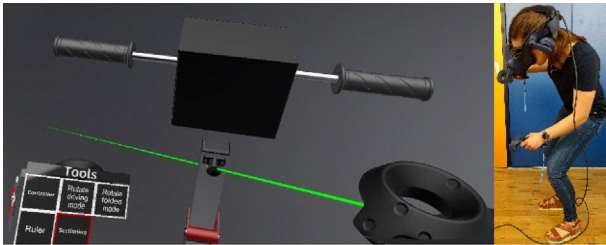


Fig. 6 Participant identifying the mechanism for adjusting the steering bar height in the IVE.

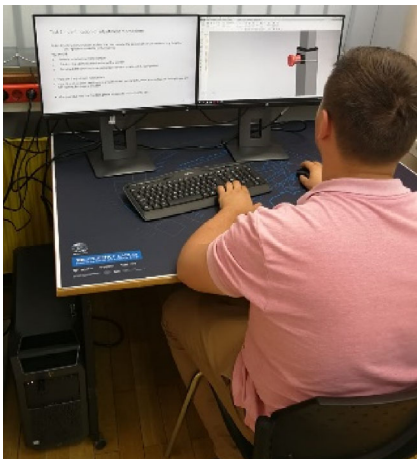


Fig. 7 Participant identifying the mechanism for adjusting the steering bar height in the nIVE.

The number of correctly identified mechanisms (N_{cam}) in the IVE was 3.3 ± 0.7 (Mean value \pm Stdev), $M: 3, R: 2-5$. That in the nIVE was 3.4 ± 0.9 , $M: 3, R: 2-5$. As shown in Fig. 8, the majority of the participants in both VEs correctly identified the mechanisms for height adjustment (seat, steering bar, and armrest height), whereas only two participants in every VE recognized a mechanism for adjusting seat angle. The Wilcoxon test did not imply a statistically significant difference in the number of correctly identified mechanisms in the IVE and nIVE ($p = 0.5369$).

Figure 9 presents the number of correctly identified

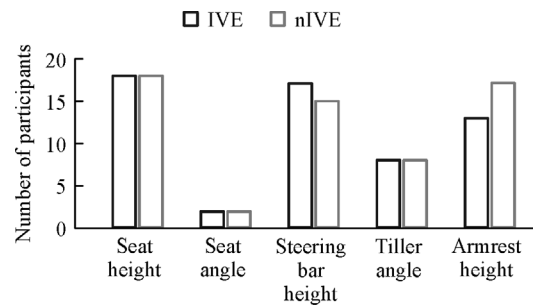


Fig. 8 Correctly identified adjustment mechanisms.

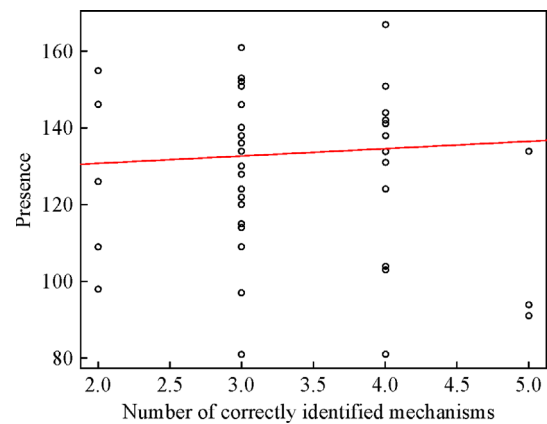


Fig. 9 Linear regression of presence–adjustment mechanisms.

mechanisms plotted against the presence that participants felt in the VE. ANOVA did not confirm the statistical validity ($p = 0.559$, $F = 0.348$) of the linear regression ($R^2 = 0.001007$, $p = 0.3982$, $F_{stat} = 1.013$), with a high number of correctly identified mechanisms when participants felt highly present in the VE.

The number of incorrectly identified mechanisms (N_{iam}) was defined as the difference between the number of correctly identified mechanisms and the total number of identified mechanisms to compare the number of errors that participants made in the VEs. It was lower in the nIVE (0.9 ± 0.6 , $M: 1, R: 0-4$) than in the IVE (1.6 ± 1.1 , $M: 1, R: 0-2$). The Wilcoxon test showed that the difference between the IVE and nIVE was significant ($p = 0.01859$). With a statistical confidence level of 95%, the number of incorrectly identified mechanisms was lower in the nIVE than in the IVE ($p = 0.009297$).

The analysis of the judgment of confidence of the participants toward the correctness of their answers showed that the confidence was higher in the IVE ($78.3\% \pm 12.4\%$, $M: 60\%, R: 50\%-100\%$) than in the nIVE ($71.5\% \pm 22\%$, $M: 60\%, R: 15\%-100\%$). Nevertheless, the correctness of answers was lower in the IVE ($65\% \pm 14\%$, $M: 80\%, R: 40\%-100\%$) than in the nIVE ($68\% \pm 17.2\%$, $M: 80\%, R: 40\%-100\%$). The Wilcoxon test showed that the difference in confidence was statistically insignificant. The linear regression presented

in Fig. 10 shows the positive correlation between the confidence judgment and the presence felt in the VE. On this basis, a participant with high experienced presence could be highly confident in his or her answer, regardless of the correctness of the answer.

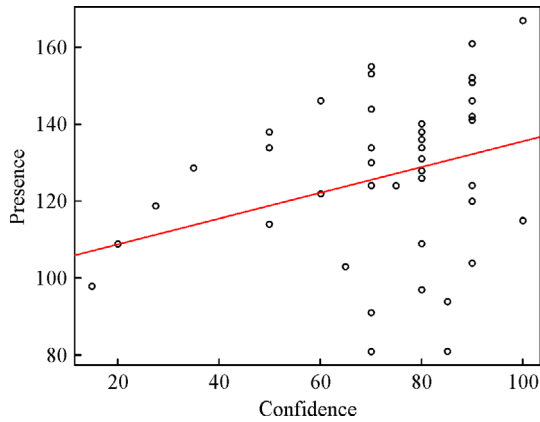


Fig. 10 Linear regression of presence–confidence.

Table 3 contains the mean values and standard deviations of the number of correctly identified adjustment mechanisms for a specific aspect of a prior experience (educational background, year of study, number of conducted DRs before the experiment, CAD skills, number of seat adjustments, and number of steering bar adjustments). The Kruskal–Wallis test showed a significant difference in the number of correctly identified adjustment mechanisms based on educational background ($p = 0.04503$), that is, ME (3 ± 0.5), ID (3.4 ± 0.7), SEE (2 ± 0), and PD (3.6 ± 0.8). No significant difference was found for the year of study ($p = 0.222$), CAD skills ($p = 0.269$), number of conducted DRs ($p = 0.06922$), number of seat adjustments ($p = 0.222$), or number of steering bar adjustments ($p = 0.8093$).

4.2 Folding steps

In task 2, the participants were asked to identify the folding steps of the scooter, which were needed to transform it from the driving mode to the folded mode. The transformation had 6 folding steps, namely, backrest, seat, seat holder, tiller, middle (Figs. 11 and 12), and swivel seat rotations.



Fig. 11 Participant identifying a folding step in the IVE.

The number of correctly identified steps (N_{cfs}) in the IVE was 3.8 ± 1.2 , $M: 4$, $R: 1-5$. In the nIVE, the number was 4.5 ± 0.9 , $M: 5$, $R: 2-5$. Figure 13 presents the number of participants who correctly identified each folding step. The tiller and middle rotations were identified by the largest number of participants in both VEs (19 in the nIVE and 17 in the IVE), whereas only one participant in each VE recognized a swivel seat rotation of 180° . The Wilcoxon test confirmed a statistically significant difference between the IVE and nIVE ($p = 0.04388$). With a statistical confidence level of 95%, the number of correctly identified folding steps was higher in the nIVE than in the IVE.

As shown in Fig. 14, a linear regression was used to analyze the relationship between the number of correctly identified folding steps and the presence experienced in the

Table 3 Relationship between the number of correctly identified adjustment mechanisms and prior experience aspects

Educational background	ME	ID	SEE	PD	
N_{cam}	3.0 ± 0.5	3.4 ± 0.7	2.0 ± 0.0	3.6 ± 0.8	
Year of study	1	2	3	4	5
N_{cam}	2.0 ± 0.0	2.8 ± 0.4	3.5 ± 0.7	3.4 ± 0.8	3.3 ± 0.8
Number of DRs	0	1–3	4–6	7–9	≥ 10
N_{cam}	3.1 ± 0.6	2.9 ± 0.5	4.2 ± 0.7	1.7 ± 0.5	3.0 ± 1.0
CAD skills	None	Limited	Basic	Intermediate	Advanced
N_{cam}	2.0 ± 0.5	3.0 ± 0.8	3.2 ± 0.5	3.5 ± 0.8	3.8 ± 0.8
Seat adjustments	0	1–5	6–10	11–30	≥ 30
N_{cam}	-	2.8 ± 0.4	3.4 ± 0.5	3.1 ± 0.9	3.5 ± 0.8
Steering bar adjustments	0	1–5	6–10	11–30	≥ 30
N_{cam}	3.1 ± 0.3	3.4 ± 0.9	2.7 ± 1.0	2.9 ± 1.4	3.4 ± 0.7

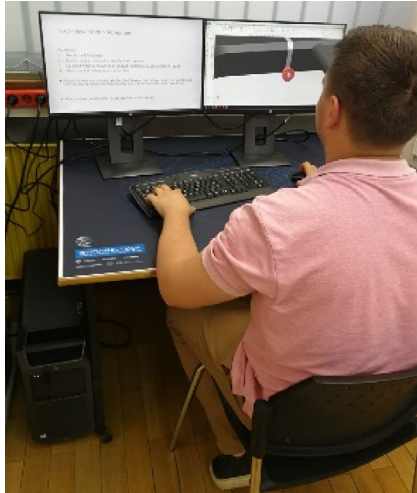


Fig. 12 Participant identifying a folding step in the nIVE.

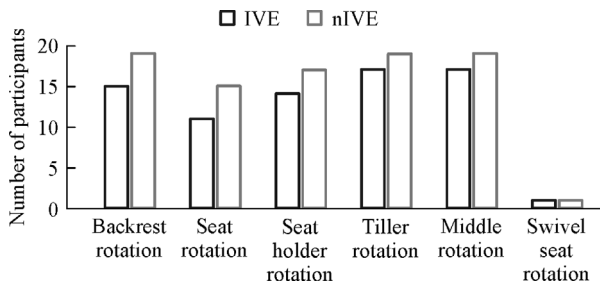


Fig. 13 Correctly identified folding steps.

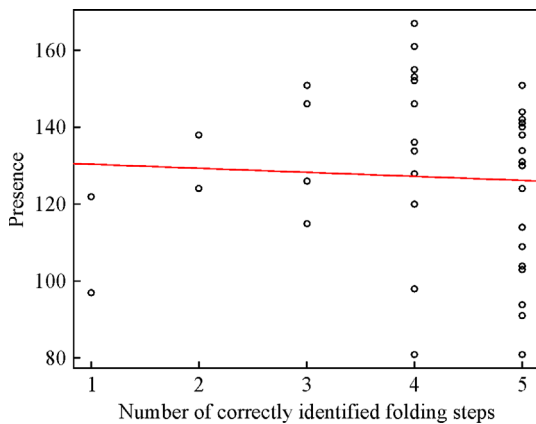


Fig. 14 Linear regression of presence–folding steps.

VE. ANOVA did not confirm the statistical validity ($p = 0.0641$, $F = 3.636$) of the linear regression model ($R^2 = 0.03494$, $p = 0.2701$, $F_{\text{stat}} = 1.353$). The regression line showed an incline in the number of correctly identified folding steps with low experienced presence, and the negative trend of the line indicated that high experienced presence did not unambiguously lead to high performance. Most of the participants (11) with experienced presence

lower than 130 correctly identified 5 folding steps, and all of them conducted the DR in the nIVE.

The participants were more confident in the correctness of their answers in the nIVE ($75.8\% \pm 13\%$, $M: 75\%$, $R: 35\%–100\%$) than in the IVE ($64\% \pm 24\%$, $M: 65\%$, $R: 0\%–90\%$). A correct answer meant the scooter was transformed from the driving mode to the folded mode by conducting the identified steps. Nevertheless, the Wilcoxon test showed that the difference in confidence was statistically insignificant ($p = 0.1483$). Figure 15 shows the results of a linear regression between confidence and experienced presence ($R^2 = 0.2816$, $p = 0.03728$, $F_{\text{stat}} = 2.274$). Most of the participants (92%) whose confidence was higher than 80% also had high experiences of presence (above 130).

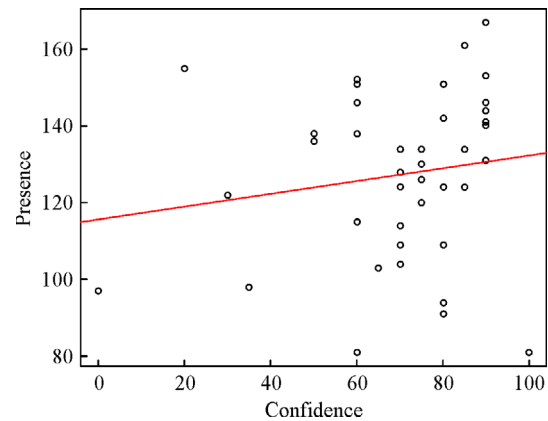


Fig. 15 Linear regression of presence–confidence.

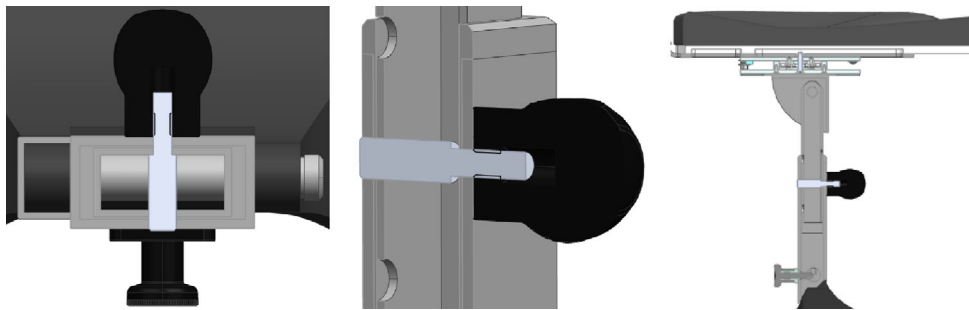
Table 4 indicates the mean values and standard deviations of the number of correctly identified folding steps for a specific aspect of prior experience (educational background, year of study, number of conducted DRs before the experiment, CAD skills, or number of vehicles folded). The Kruskal–Wallis test did not show a significant correlation between task performance and year of study ($p = 0.1289$), educational background ($p = 0.6113$), CAD skill ($p = 0.2019$), number of previously conducted DRs ($p = 0.2407$), or number of vehicles folded ($p = 0.5035$).

4.3 Detection of functional errors

In task 3, the participants were asked to inspect the seat height mechanism for functional errors. To detect a functional error, they first had to understand the function and working principle of the mechanism. As presented in Fig. 16, the analyzed mechanism had the following three functional errors: 1) The pin could not be pulled out because its diameter was wider than the diameter of the hole; 2) Two holes were missing in the outer profile; and 3) The seat could not be adjusted to the lowest position because the vertical dimension of the seat rotation element was excessively high.

Table 4 Relationship between the number of correctly identified folding steps and prior experience aspects

Educational background	ME	ID	SEE	PD	
N_{cfs}	4.0 ± 1.3	4.3 ± 1.0	3.5 ± 0.5	4.1 ± 1.1	
Year of study	1	2	3	4	5
N_{cfs}	4.0 ± 0.0	3.0 ± 0.7	4.3 ± 1.0	4.1 ± 1.3	4.7 ± 0.5
Number of DRs	0	1–3	4–6	7–9	≥ 10
N_{cfs}	3.7 ± 1.3	4.6 ± 0.5	4.6 ± 0.5	4.5 ± 0.5	4.6 ± 0.8
CAD skills	None	Limited	Basic	Intermediate	Advanced
N_{cfs}	2.3 ± 1.4	4.0 ± 0.8	3.9 ± 1.1	4.4 ± 1.0	4.8 ± 0.4
Number of vehicles folded	0	1–5	6–10	11–30	≥ 30
N_{cfs}	3.3 ± 2.4	3.7 ± 1.3	4.2 ± 1.1	4.4 ± 0.5	3.6 ± 1.3

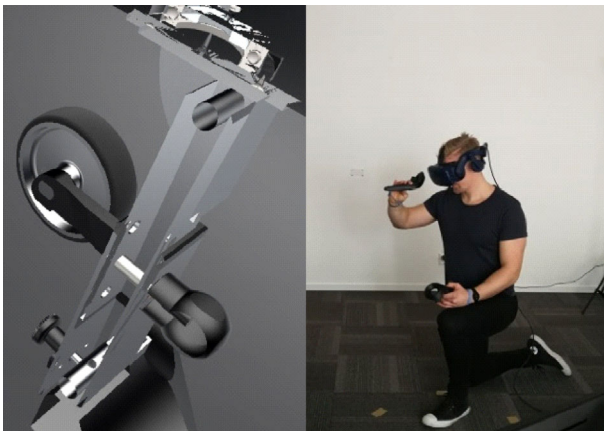
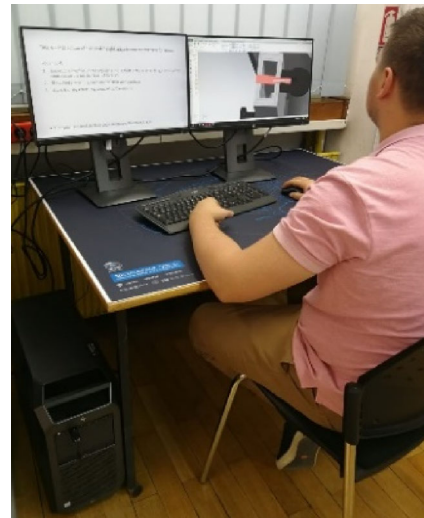
**Fig. 16** Functional errors (left—Error 1, middle—Error 2, right—Error 3).

In the IVE (Fig. 17), seven participants detected one functional error—three detected Error 1 (confidence: 80%), three detected Error 2 (confidence: 20%, 50%, 70%), and one detected Error 3 (confidence: 30%). In the nIVE (Fig. 18), six participants detected only Error 1 (confidence: 50% (2), 60% (2), and 70% (2)), and one participant detected the two other errors. The confidence of the participants who did not detect any functional error that the mechanism would work was $44\% \pm 24.8\%$, $M: 50\%$, $R: 10\%–90\%$ in the IVE and $49\% \pm 26.7\%$, $M: 65\%$, $R: 0\%–80\%$ in the nIVE. Statistical analysis of the prior

experience and presence of participants was not conducted due to the small sample size.

5 Discussion

This study explored the effects of IVE and nIVE on conceptual DR activities. Specifically, the study focused on

**Fig. 17** Participant detecting an error in the IVE.**Fig. 18** Participant detecting an error in the nIVE.

the ability of engineering students to identify the mechanisms of a technical system and understand their functions. Such ability was investigated through three tasks defined based on the product requirements stated by the manufacturing company (an industrial partner in the project-based course during which the reviewed design solution was developed).

5.1 Relationship between DR task performance and VEs

Understanding mechanisms and their functions is one of the core activities of DRs, and decisions on further steps in design projects often rely on it (Dieter and Schmidt, 2012; Freeman et al., 2016). On the basis of the work of Newcombe and Shipley (2015), DR tasks, through which the ability of engineering students to identify mechanisms and understand their functions was investigated, can be associated with the spatial skills needed for performance. The internal static skill of an individual is related to the identification and understanding of the spatial features of objects (the scooter), including their size and the arrangement of their parts (parts and subassemblies of the scooter). It can also be associated with the first task (identification of adjustment mechanisms and understanding their working principles) and the third task (detection of functional errors). The spatial skill associated with the identification of the folding steps that should be followed to transform the scooter from driving mode to folded mode is an intrinsic dynamic skill. It enables participants to mentally transform 2D or 3D objects (parts and subassemblies of the scooter) and visualize the design solution as a series of transformations, such as rotations and folding (Newcombe and Shipley, 2015). As hypothesized in Section 1, the ability of engineering students to identify mechanisms and understand their functions was expected to be enhanced in an IVE due to the benefits and positive effects of stereoscopic display on their visual perception, spatial skills, and spatial comprehension (as discussed in Section 2). Nevertheless, the results of the statistical analysis showed no statistically significant difference in the number of correctly identified adjustment mechanisms in the IVE and nIVE or the confidence of participants in the correctness of their answers. The number of incorrectly identified mechanisms was significantly lower in the nIVE than in the IVE. The number of correctly identified folding steps was significantly higher in the nIVE ($M = 5$) than in the IVE ($M = 4$).

Design error detection is a common DR activity because the early detection of errors may diminish PD costs and enhance user satisfaction (de Casenave and Lugo, 2017). Several scholars (Dunston et al., 2010; Satter and Butler, 2012; Liu et al., 2014; Berg and Vance, 2017) reported results showing that DR in an IVE often reveals unexpected design problems or errors, which are usually undetected when a design is reviewed in an nIVE. In Wolfartsberger (2019)'s study, the participants were more

successful with finding errors using IVR than using the conventional user interface, but the difference was small. Aligned with these findings, the number of participants in this experimental study who detected at least one functional error of the seat height mechanism was higher in the IVE (7) than in the nIVE (6). However, the observed difference was small.

The results of this experiment showed either an insignificant difference in the identification of mechanisms and the understanding of their functions among VEs or higher performance in the nIVE. This finding led to the conclusion that stereoscopic display with a first-person view and navigation inside the IVE did not improve the ability of engineering students (who participated in the experiment) to identify mechanisms and understand their functions. Hypothesis (1) was rejected, and the results were not in line with the generic assumption (derived from the literature) that reviewing a 3D CAD model in an IVE improves the understanding of users regarding how components are laid out spatially and how each component interacts with others (Berg and Vance, 2017).

The results cannot be unambiguously explained, but some of the possible explanations mentioned in the literature are similar, that is the visual information provided by both VEs (Hannah et al., 2012), the complexity of the reviewed design solution (de Casenave and Lugo, 2017), the misinterpretation of participants regarding a design, and a high effect of prior experience rather than visualization technology (Banerjee et al., 2002; Thimmaiah et al., 2017). With regard to the last item, a statistically significant correlation was found in this study between the educational background of participants and their identification and understanding of adjustment mechanisms ($p = 0.04503$). PD students correctly identified the highest number of adjustment mechanisms (3.6 ± 0.8), followed by ID students (3.4 ± 0.7). In contrast, a statistically significant correlation was not found between the same aspect of prior experience and the number of correctly identified folding steps. The order of the tasks might have affected the results. This condition might imply that educational background had a great effect on the results when participants were inconsiderably familiar with the presented design solution. These results raise a question on the relationship between spatial skills and educational background, which should be further investigated. The results did not show a consistent relationship with educational background for all tasks. Differences in educational background were noticeable for tasks associated with the static spatial skill. Previous studies (Freeman et al., 2016; de Casenave and Lugo, 2017) reported that the different complexity levels of reviewed technical systems might affect the understanding of participants about VE. Although no significant difference in the understanding of subassembly functions was observed when a technical system of the third complexity level (Hubka and Eder, 1988) was reviewed,

the results could be different for technical systems of other complexity levels.

5.2 Relation between DR task performance and experienced presence

In this study, insights into the presence that engineering students experienced in VEs were gained through their completion of the PQ (Witmer and Singer, 1998). Although unusual, high presence in nIVE and low presence in IVE are possible (Schubert et al., 2001; Banerjee et al., 2002) when considering the subjective experience of users rather than the technology used (Slater et al., 1995; Witmer and Singer, 1998; Faas et al., 2014). Nevertheless, the results of this study showed that engineering students who reviewed the scooter in IVE rated their experienced presence higher (140.15 ± 15.53 , $M: 140$, $R: 97\text{--}167$) than those who reviewed it in nIVE (116.75 ± 17.52 , $M: 117.5$, $R: 81\text{--}144$).

As discussed in Section 2, performance in a VE is related to the presence that participants experience, although this relationship is not fully defined in the literature (Hendrix and Barfield, 1996; Hubona et al., 1997; Banerjee et al., 2002; Vora et al., 2001; Paes et al., 2017). In this study, the relationship between the presence experienced in VEs and the ability of engineering students to identify mechanisms and understand their functions was analyzed via regression analysis (Figs. 9 and 14). However, ANOVA did not confirm the statistical validity of the model. Most participants (61%) who correctly identified 5 folding steps completed the PQ with a sum less than 130 (while $M: 128.5$). In contrast, the experienced presence of the majority (60%) of participants who correctly identified four or five adjustment mechanisms was scored higher than 130. These results were in line with the findings of Faas et al. (2014), which showed a correlation between high presence and high- and low-quality designed solutions, whereas low presence was correlated with average-quality designed solutions. Other studies (Hubona et al., 1997; Witmer and Singer, 1998; Banerjee et al., 2002) reported a correlation between high presence and task performance in VEs, but it was not significant in all cases. Scholars explained those results through a noticeable effect of prior experience and the subjective interpretation of design, tasks, and environments. For example, Banerjee et al. (2002) found an insignificant correlation between presence and design-understanding scores in their experiment. This finding could be explained by low experience of participants with using IVR technology. In the study, a significant correlation among the experiences of participants with IVR technology was not observed because most of the participants were equally inexperienced, as presented in Table 2. Nevertheless, the two participants with the highest experience had the best performance (at least four correctly

identified mechanisms in both tasks and correctly detected functional error). One participant had the highest experience in using IVR technology (780 min). The other participant had low experience using IVR technology but had the highest experience in the number of conducted DRs before the experiment (4) and years of experience in engineering (3).

5.3 Confidence judgment

In general, humans intuitively evaluate their answers concerning conducted tasks, even if feedback about their performance is unavailable (Spence et al., 2016). An investigation of confidence judgment opens a discussion on the awareness of participants of their design comprehension and has implications for shared understanding during DR (Eckert et al., 2012). Confidence may reflect the strength of accumulated sensory evidence at the time of conducting a task (Spence et al., 2016). The brain estimates the reliability of encoded information and provides feelings of confidence (Spence et al., 2016). Linear regression analyses, in accordance with Figs. 10 and 15, showed a positive correlation between confidence and experienced presence in VEs. A participant with high experienced presence was considerably confident with his/her answer regarding the understanding of mechanisms. For example, the average level of confidence in the correctness of the identified folding steps for participants with experienced presence lower than 130 was 66.05 ± 22.04 . For those with experienced presence higher than 130, the confidence level was 73.33 ± 18.21 . Similarly, for the level of confidence in the correctness of identified adjustment mechanisms, the average result was 71.11 ± 22.33 for those with experienced presence lower than 130 and 78.1 ± 13.32 for those with experienced presence higher than 130. ANOVA did not confirm the validity of the linear regression model between confidence and the number of correctly identified adjustment mechanisms ($p = 0.251$, $F = 1.361$) or folding steps ($p = 0.0722$, $F = 3.419$), but they were described with a positive trend line in both cases. These results were in accordance with the results of a study performed by Hannah et al. (2012), which indicated a high level of confidence when looking at a high level of fidelity representations. The correctness of the answers of participants was higher in the nIVE for adjustment mechanisms and folding step identification. However, the participants were more confident in the correctness of adjustment mechanism identification in the IVE, than that of the folding step identification in the nIVE. These findings supported the conclusions of a previously conducted study (Hannah et al., 2012), indicating that although participants might be more confident in their answers in one VE than in the other, they were not necessarily more correct.

5.4 Limitations of the study

The first limitation of this study concerns the low practical engineering experience of participants. The results should be related to engineering students, not engineers in general. IVR may improve the identification of mechanisms and the understanding of their functions for other groups of examinees. Therefore, the insights and conclusions obtained are restricted to the engineering students who have participated in this experiment. In addition, the experiment has been conducted at two universities by two researchers. Although the setup and experimental procedure are the same, the different locations may have affected the experimental results.

6 Conclusions and further research

A conceptual DR experiment was conducted in this study to explore differences in the ability of engineering students to identify the mechanisms of a technical system and understand their functions when the design solution was presented in an IVE (using IVR technology) and in an nIVE (using a conventional user interface). This ability was measured through three DR tasks and associated questions regarding judgment of confidence of participants in the correctness of their answers. In contrast to Hypothesis (1), IVR technology did not support the enhanced ability of engineering students to identify mechanisms and understand their functions. The results can be partially explained by the prior experience of participants having a higher effect than visualization technology had. Linear regression revealed positive (adjustment mechanisms) and negative (folding steps) trend lines between experienced presence and DR task performance. However, the models were statistically insignificant. Therefore, the results did not provide an answer to Hypothesis (2) and further research was needed to form conclusions on the relationship between experienced presence and an understanding of a subassembly function. The findings concerning confidence judgment indicated that high experienced presence positively affected the confidence of engineering students in their understanding of the mechanism function. Nevertheless, this study only covered the surface of the confidence aspect, its relationship with visualization technologies, and the implications for shared understanding. One avenue for further work included the repetition of the experiment with engineers of various expertise levels. Although the number of study participants seemed to be appropriate compared with that in the experiments of other scholars, the sample could be larger in future studies. As discussed in Section 5.2, the complexity level of the reviewed technical system might affect the results of the study. Therefore, the understanding of engineers regarding technical systems of

various complexity levels must be researched and compared.

References

- Aust M, de Clerk M, Blach R, Dangelmaier M (2011). Towards a holistic workflow pattern for using VR for design decisions: Learning from other disciplines. In: ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers Digital Collection, 1463–1470
- Banerjee P, Bochenek G M, Ragusa J M (2002). Analyzing the relationship of presence and immersive tendencies on the conceptual design review process. *Journal of Computing and Information Science in Engineering*, 2(1): 59–64
- Bassanino M, Wu K C, Yao J, Khosrowshahi F, Fernando T, Skjærbæk J (2010). The impact of immersive virtual reality on visualisation for a design review in construction. In: Proceedings of the 14th International Conference on Information Visualisation. IEEE, 585–589
- Berg L P, Vance J M (2017). An industry case study: Investigating early design decision making in virtual reality. *Journal of Computing and Information Science in Engineering*, 17(1): 011001
- Chandrasegaran S K, Ramani K, Sriram R D, Horváth I, Bernard A, Harik R F, Gao W (2013). The evolution, challenges, and future of knowledge representation in product design systems. *Computer-Aided Design*, 45(2): 204–228
- Chen Q, Weidner R, Vossel S, Weiss P H, Fink G R (2012). Neural mechanisms of attentional reorienting in three-dimensional space. *Journal of Neuroscience*, 32(39): 13352–13362
- Chu C C P, Dani T H, Gadh R (1998). Evaluation of virtual reality interface for product shape designs. *IIE Transactions*, 30(7): 629–643
- Coburn J Q, Freeman I, Salmon J L (2017). A review of the capabilities of current low-cost virtual reality technology and its potential to enhance the design process. *Journal of Computing and Information Science in Engineering*, 17(3): 031013
- Cruz-Neira C, Sandin D J, DeFanti T A (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, 135–142
- de Casenave L, Lugo J E (2017). Design review using virtual reality enabled CAD. In: Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Cleveland: American Society of Mechanical Engineers Digital Collection, 1–11
- Dieter G E, Schmidt L C (2012). *Engineering Design*. 5th ed. New York: The McGraw-Hill Higher Education
- Dunston P S, Arns L L, McGlothlin J D (2010). Virtual reality mock-ups for healthcare facility design and a model for technology hub collaboration. *Journal of Building Performance Simulation*, 3(3): 185–195
- Eckert C, Ruckpaul A, Alink T, Albers A (2012). Variations in functional decomposition for an existing product: Experimental results. *Artificial Intelligence for Engineering Design, Analysis and*

- Manufacturing, 26(2): 107–128
- Faas D, Bao Q, Frey D D, Yang M C (2014). The influence of immersion and presence in early stage engineering designing and building. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 28(2): 139–151
- Freeman I J, Salmon J L, Coburn J Q (2016). CAD integration in virtual reality design reviews for improved engineering model interaction. In: *ASME International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers Digital Collection
- George P, Kemeny A, Colombet F, Merienne F, Chardonnet J R, Thouvenin I M (2014). Evaluation of smartphone-based interaction techniques in a CAVE in the context of immersive digital project review. In: *The Engineering Reality of Virtual Reality*. International Society for Optics and Photonics, 9012: 901203
- Germani M, Mengoni M, Peruzzini M (2009). Metric-based approach for VR technology evaluation in styling product design. In: *Proceedings of the ASME 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 1325–1339
- Hannah R, Joshi S, Summers J D (2012). A user study of interpretability of engineering design representations. *Journal of Engineering Design*, 23(6): 443–468
- Hartmann T, Wirth W, Schramm H, Klimmt C, Vorderer P A, Gysbers A, Böcking S, Ravaja N, Laari J, Saari T, Gouveia F R, Sacau A (2015). The Spatial Presence Experience Scale (SPES): A short self-report measure for diverse media settings. *Journal of Media Psychology*, 28(1): 1–15
- Hendrix C, Barfield W (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators & Virtual Environments*, 5(3): 274–289
- Heydarian A, Carneiro J P, Gerber D, Becerik-Gerber B, Hayes T, Wood W (2015). Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations. *Automation in Construction*, 54: 116–126
- Hubka V, Eder W E (1988). *Theory of Technical Systems: A Total Concept Theory for Engineering Design*. Berlin: Springer-Verlag
- Hubona G S, Shirah G W, Fout D G (1997). The effects of motion and stereopsis on three-dimensional visualization. *International Journal of Human-Computer Studies*, 47(5): 609–627
- Igroup (2019). *Igroup Presence Questionnaire (IPQ)*
- Kalisperis L N, Muramoto K, Balakrishnan B, Nikolic D, Zikic N (2002). Evaluating relative impact of virtual reality system variables on architectural design comprehension and presence: A variable-centered approach using fractional factorial experiment. In: *Proceedings of the 24th Education and Research in Computer Aided Architectural Design in Europe (eCAADe)*. Volos, Greece, 66–73
- Lessiter J, Freeman J, Keogh E, Davidoff J (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators & Virtual Environments*, 10(3): 282–297
- Liu Y F, Lather J, Messner J (2014). Virtual reality to support the integrated design process: A retrofit case study. In: *International Conference on Computing in Civil and Building Engineering*, 801–808
- Lombard M, Ditton T (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2): JCMC321
- Lombard M, Ditton T B, Weinstein L (2009). Measuring presence: The temple presence inventory. In: *Proceedings of the 12th Annual International Workshop on Presence*. Los Angeles, 1–15
- Maftei L, Harty C (2016). Challenging design perceptions in immersive virtual reality environments? In: *Proceedings of the 32nd Annual ARCOM Conference*, 113–122
- Mckenna A F, Chen W, Simpson T W (2008). Exploring the impact of virtual and physical dissection activities on students' understanding of engineering design principles. In: *Proceedings of the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 359–368
- Mengoni M, Peruzzini M, Germani M (2009). The impact of virtual environments on human collaboration in product design. In: *Proceedings of the 17th International Conference on Engineering Design—vol. 9, Human Behavior in Design*. Palo Alto, CA, 57–68
- Newcombe N S, Shipley T F (2015). Thinking about spatial thinking: New typology, new assessments. In: Gero J S, ed. *Studying Visual and Spatial Reasoning for Design Creativity*. Dordrecht: Springer, 179–192
- Ostergaard K J, Wetmore W, Summers J D (2003). A methodology for the study of the effects of communication method on design review effectiveness. In: *Proceedings of ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 383–390
- Paes D, Arantes E, Irizarry J (2017). Immersive environment for improving the understanding of architectural 3D models: Comparing user spatial perception between immersive and traditional virtual reality systems. *Automation in Construction*, 84: 292–303
- Pahl G, Beitz W, Feldhusen J, Grote K H (2007). *Engineering Design: A Systematic Approach* (3rd ed). London: Springer-Verlag
- Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition*, 28(1): 39–58
- Rigutti S, Stragà M, Jez M, Baldassi G, Carnaghi A, Miceu P, Fantoni C (2018). Don't worry, be active: How to facilitate the detection of errors in immersive virtual environments. *PeerJ*, 6: e5844
- Robinson M A (2016). Quantitative research principles and methods for human-focused research in engineering design. In: Cash P, Stanković T, Štorga M, eds. *Experimental Design Research: Exploring Complex Design Activity*. Cham, Switzerland: Springer, 41–64
- Satter K, Butler A (2015). Competitive usability analysis of immersive virtual environments in engineering design review. *Journal of Computing and Information Science in Engineering*, 15(3): 031001
- Satter K M, Butler A C (2012). Finding the value of immersive, virtual environments using competitive usability analysis. *Journal of Computing and Information Science in Engineering*, 12(2): 024504
- Schnabel M A, Kvan T (2003). Spatial understanding in immersive virtual environments. *International Journal of Architectural Computing*, 1(4): 435–448

- Schubert T, Friedmann F, Regenbrecht H (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3): 266–281
- Slater M, Linakis V, Usoh M, Kooper R (1995). Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In: *Proceedings of ACM Symposium on Virtual Reality Software and Technology (VRST)*. New York: ACM Press, 163–172
- Spence M L, Dux P E, Arnold D H (2016). Computations underlying confidence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, 42(5): 671–682
- Stankov L, Kleitman S, Jackson S A (2014). Measures of the trait of confidence. In: Boyle G J, Saklofske D H, Matthews G, eds. *Measures of Personality and Social Psychological Constructs*. Academic Press, 158–189
- Stanney K M, Mourant R R, Kennedy R S (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators & Virtual Environments*, 7(4): 327–351
- Steuer J (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4): 73–93
- Sutherland I E (1968). A head-mounted three dimensional display. In: *Proceedings of the Fall Joint Computer Conference, part I*, 757–764
- Thimmaiah S, Phelan K, Summers J D (2017). An experimental study on the influence that failure number, specialization, and controls have on confidence in predicting system failures. *Journal of Mechanical Design*, 139(1): 011102
- Vora J, Nair S, Gramopadhye A K, Melloy B J, Medlin E, Duchowski A T, Kanki B G (2001). Using virtual reality technology to improve aircraft inspection performance: Presence and performance measurement studies. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Los Angeles, CA: SAGE Publications, 45 (27): 1867–1871
- Wann J, Mon-Williams M (1996). What does virtual reality NEED? Human factors issues in the design of three-dimensional computer environments. *International Journal of Human-Computer Studies*, 44 (6): 829–847
- Watanuki K (2010). Development of virtual reality-based universal design review system. *Journal of Mechanical Science and Technology*, 24(1): 257–262
- Whyte J, Bouchlaghem D, Thorpe T (2002). IT implementation in the construction organization. *Engineering, Construction and Architectural Management*, 9(5–6): 371–377
- Witmer B G, Singer M J (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators & Virtual Environments*, 7(3): 225–240
- Wolfartsberger J (2019). Analyzing the potential of virtual reality for engineering design review. *Automation in Construction*, 104: 27–37