Susanna Aromaa

VTT Technical Research Centre of Finland
Tel. +358 40 724 9828
susanna.aromaa@vtt.fi

Kaisa Väänänen

Tampere University of Technology
kaisa.vaananen@tuni.fi

Suitability of virtual prototypes to support human factors/ergonomics evaluation during the design

Abstract

In recent years, the use of virtual prototyping has increased in product development processes, especially in the assessment of complex systems targeted at end-users. The purpose of this study was to evaluate the suitability of virtual prototyping to support human factors/ergonomics evaluation (HFE) during the design phase. Two different virtual prototypes were used: augmented reality (AR) and virtual environment (VE) prototypes of a maintenance platform of a rock crushing machine. Nineteen designers and other stakeholders were asked to assess the suitability of the prototype for HFE evaluation. Results indicate that the system model characteristics and user interface affect the experienced suitability. The VE system was valued as being more suitable to support the assessment of visibility, reach, and the use of tools than the AR system. The findings of this study can be used as a guidance for the implementing virtual prototypes in the product development process.

Keywords (max 3): virtual prototyping, human factors/ergonomics, virtual environment

1. Introduction

A virtual prototype is a computer simulation of a physical product that can be presented, analysed and tested from various aspects. The process of constructing and testing a virtual prototype is called virtual prototyping (VP) (Wang, 2002). In recent years, the use of VP has increased in the product development process due to the improved availability and lowered prices of VP technologies (Choi et al., 2015). However, companies do not necessary know how to use VP technologies effectively, and for that reason they do not gain the full potential from it.

Virtual prototyping supports the evaluation of human factors/ergonomics (HFE) already in the early design phase. According to the principles of human-centred design (HCD) ISO 9241-210 (2010) and participatory design (Muller and Kuhn, 1993) of interactive systems, it is crucial to involve end-users and other stakeholders in the design and evaluation of technological products. International Ergonomics Association (IEA, 2000) defines HFE as "the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance". Similarly, "Practitioners of ergonomics and ergonomists contribute to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people" (IEA, 2000). According to Dul et al. (2012), HFE seeks to improve performance and well-being through systems design.

Virtual prototypes can be different in their level of virtuality and fidelity. Milgram et al. (1995) have developed a reality–virtuality continuum which is a continuous scale ranging between the completely virtual, virtuality, and the completely real, reality. Using the definition by Kalawsky (1993), virtual environment (VE) uses virtual reality (VR) technologies in order to provide human beings with the means of manipulation and sensory modalities. In practice, it means that humans are able to navigate in the VE (e.g. move from one place to another), manipulate objects (e.g. turn a steering wheel) and get sensory feedback (e.g. visual or auditory). The term mixed reality describes environments between virtual and real. An example of mixed reality is augmented reality (AR), which means that the user is able to see the real world, with virtual objects superimposed upon or composited with the real world (Azuma, 1997).

Several studies (e.g. Bordegoni et al., 2009; Bullinger and Dangelmaier, 2003; Cecil and Kanchanapiboon, 2007; Karkee et al., 2011; Kremer, 1998; Kim et al., 2011; Kulkarni et al., 2011; Lawson et al., 2016; Park et al., 2009; Seth et al., 2011) state that VP has been considered as a powerful prototyping solution to overcome the shortcomings of conventional prototyping methods. They conclude that the production of a physical prototype is costly and time-consuming and, therefore, the reduction of the number of physical prototypes would shorten the time to market. Mujber et al. (2004) summarise the benefits of virtual reality in manufacturing applications in three categories: design, operations management and manufacturing processes. The benefits at technological, design and business levels are described by Aromaa et al. (2012) . In addition, Leino (2015) models the business and organisational value of VP.

In the prototype fidelity domain, there are related studies that do not apply virtual reality techniques but compare, for example, computer and paper prototypes (Boothe et al., 2013; Lim et al., 2006; Sauer and Sonderegger, 2009; Sauer et al., 2010). These studies show that the main usability issues can be identified with prototypes of different fidelity levels. Some usability issues, however, cannot be evaluated using these prototypes, and therefore, Lim et al. (2006) state that it is important to determine what aspects need to be evaluated before building low-fidelity prototypes.

Perez and Neumann (2015) requested consideration of VP tools in supporting the integration of HFE issues in the design of new workplaces. They identified the importance of the utility of the VP tools from the ergonomists' and engineers' points of view, also listing categories to be considered, such as time, cost, training, difficulty to use, trustworthiness, graphics, flexibility, usefulness, and report presentation. Other approaches to support the development and usability of VP systems have been suggested by Stanney et al. (2003); Sutcliffe and Gault (2004); Eastgate et al. (2014). In addition, Jia et al. (2012) proposed a method for the design of more usable and efficient virtual training systems. Canuto da Silva and Kaminski (2015) proposed a procedure for the selection of virtual and physical prototypes in the product development process.

According to Ma et al. (Ma et al., 2011), the collaborative VE is a useful tool for supporting complex product design. Therefore, VP can be used to support communication and interaction between different stakeholders during design reviews (Aromaa et al., 2012; Bordegoni et al., 2009; Bordegoni and Caruso, 2012; Kremer, 1998; Leino, 2015; Shen et al., 2010). Huet et al. (2007)claim that design reviews are efficient tools for sharing information about the product and for managing knowledge exchange. In addition, the use of VP during the HCD is a complex task and therefore approaches to support the use of virtual prototypes in HCD have been developed (Barbieri et al., 2013; Bordegoni et al., 2009; Bordegoni et al., 2014; Broberg et al., 2011; Ferrise et al., 2013; Hall-Andersen and Broberg, 2014; Mahdjoub et al., 2013).

The use of VP in HFE evaluation has been studied in several research projects such as those by Wilson and D'Cruz, 2006; Bullinger and Dangelmaier, 2003; Park et al., 2009; Bordegoni et al., 2009; Karkee et al., 2011. It seems that the fidelity of the prototype does not affect the subjective evaluation of the usability of the product, but it affects the task performance and therefore the HFE evaluations. Bruno and Muzzupappa (2010) discovered that VR techniques are valid alternatives to traditional methods for the usability evaluation of product interfaces, and that the interaction with the VE does not invalidate the usability evaluation itself. However, in VEs users may become fatigued more quickly, require more time and greater effort and experience more discomfort and more task difficulty than in a real environment (Hu et al., 2011). Therefore, Wu et al. (2012) discovered that the results from the 1991 revised NIOSH Lifting Equation RWL tool were significantly larger in a virtual prototype than in physical prototype. Pontonnier et al. (2013) compared assembly tasks in a real environment and in VEs with and without haptics. They discovered that the mechanical limitations of the haptic device lowered the sensation of presence and resulted in an increase in the difficulty compared to real environment and VEs without haptics. Lawson et al. (2015) compared virtual and physical prototypes and discovered that virtual prototypes had lower validity and reliability than physical ones for identifying entry and exit issues in passenger vehicles. Gavish et al. (2013) studied the use of VR and AR training for industrial maintenance and assembly tasks. They found that the AR system was suitable for training but the VR system's suitability needed to be evaluated further. Nee et al. (2012) review the use of AR applications in design and manufacturing.

Digital human models (DHMs) can be used for proactive analysis of HFE in design (Chaffin, 2005; Demirel and Duffy, 2007). Lämkull et al. (2009) found that DHMs have been proven to correctly predict HFE issues for standing and unconstrained working postures. In addition, DHMs can provide information to designers, for example, about workers' reach, clearance, vision, posture and strength capabilities (Feyen et al., 2000; Sanjog et al., 2015). Nevertheless, the functionality of DHMs still needs improvement (Chaffin, 2007; Lämkull et al., 2009). In this paper, however, we discuss only real users using virtual prototypes (see a mixed prototyping framework in Bordegoni et al., 2009).

Despite the research carried out in the area of VP, there is not enough knowledge of the efficient use of VP in HCD. In particular, the question regarding which type of virtual prototypes should be used in HFE evaluation remains open. Therefore, companies who use VP in design are unable to gain full potential from it. The purpose of this study was to evaluate the suitability of VP to support HFE evaluation during the design phase. Two virtual prototypes, augmented reality and virtual environment, were selected to be tested in this study. They were chosen because both technologies can be used to visualise new design solutions such as a maintenance platform for machines. The goal was to find out differences between different fidelity level prototypes in the reality–virtuality continuum. The findings of this study can provide guidance for the preparation and use of virtual prototypes in HFE evaluation. The paper is organized as follows. Section 2 presents related work. Section 3 describes the design of the study. Section 4 provides results from the tests. Section 5 discusses collected results and section 6 draws conclusions.

2. The study design

2.1 Experiment design

The goal of the study was to evaluate the suitability of VP to support HFE evaluation. A semi controlled between-group experiment was employed in the study. Nineteen participants from a company that offers minerals processing solutions and services took part in the experiment. They were designers or other stakeholders from a product lifecycle of the maintenance platform of a rock crushing machine. They all deal with HFE issues such as performance and well-being during the design process. The independent variable was the type of a virtual prototype: AR prototype and VE prototype. The two experiments will be called AR test/AR system and VE test/VE system for the remainder of this paper. Dependent variables measured in this experiment were the suitability of the virtual prototype for the HFE evaluation, and the overall assessment of the design object. In addition, subjective workload was evaluated.

2.2 Participants

Nine people from the company participated in the AR test. Six subjects were design engineers, one an assembly worker, one an assembly designer and one a project leader. All the participants were males and stakeholders from the product lifecycle and have an understanding of the maintenance task and its requirements regarding HFE issues. Their average age was 34 years (age range: 25–47 years) and their average time in their current employment was 8 years (range: 1–20 years). Their average height was 1.83 m (range: 1.70–2.03 m). Four people had never used VR technologies before, four had tried VR technologies before and one had used VR technologies more frequently.

Ten people from the company participated in the VE test. They were different from the test subjects participating in the AR test. All of them were males and design engineers. They had an understanding of the maintenance task and its requirements regarding HFE issues. Average age was 37 years (age range: 29–56 years) and their average time in their current employment was 10 years (range: 3–30 years). Their average height was 1.85 m (range: 1.74–1.98 m). Six people had never used virtual reality technologies beforehand. Four of the test subjects had tried VR technologies before.

2.3 Data collection and analysis

Questionnaires were used as data collection method in this semi-structured experiment. Participant demographics and consent forms were collected first. Next, the questionnaire about the overall assessment of a design object (a maintenance platform) was used. The questionnaire included three open questions regarding the design issues, development ideas and required information for design decision making. Participants also verified the design object on a Likert scale from one (strongly disagree) to five (strongly agree). In addition, participants explained their selection on the Likert scale. Next questionnaire was related to the suitability of the current system for HFE evaluation during the design. The questions were based on HFE checklists such as Karwowski (2006) and related to the maintenance task at hand. Participants selected on the Likert scale (one as not at all and five as very well) how well the system supported HFE evaluation. An unweighted NASA-TLX (Hart and Staveland, 1988) questionnaire was used to collect the experience of subjective workload during the use of the system. In addition, an adopted simulator sickness questionnaire (SSQ) (Kennedy et al., 1993) was used in the VE system but not in the AR system, because the AR system was not considered to be an environment that could provide SSQ symptoms in a given time frame. In applied SSQ, a four point Likert scale from none to severe was used to collet symptoms before and after the test. A content analysis was applied for qualitative data analysis and T-test in SPSS was used for the statistics.

2.4 Virtual prototypes and the test situations

The design object reviewed in the test was a maintenance platform attached to a mobile rock crushing machine. This was an upgrade for the existing machine. The purpose of the maintenance platform is to provide a safe, ergonomic and efficient workspace for maintenance workers. This study investigates the use of two different virtual prototypes (AR and VE) from the reality –virtuality continuum. The prototypes are described in more detail in the following.

In the AR test, the system includes a virtual model of the product (the maintenance platform), the real rock crushing machine, a virtual frame and a cover, a real environment, three different postures of a digital human model (DHM) and a real participant. The environment is not authentic because the machine is located outside in the backyard of the factory. The weather was bright and sunny. The temperature was around five degrees and there was a cold wind. Participants were able to stand next to the real machine but could not climb on the virtual maintenance platform (Figure 1). The models of the maintenance platform and DHMs do not have dynamic model characteristics. A user interface consists in the means of manipulation and sensory modalities. The means of manipulation were provided by buttons to show and hide models; move and scale the machine model for tracking purposes, and lock and release tracking. Different view angels could be achieved by walking around. Only visual feedback was provided as sensory modalities. Used hardware and software were iPad, Unity 3D and AR tracking with Qualcom Vuforia (Unity add-on component).

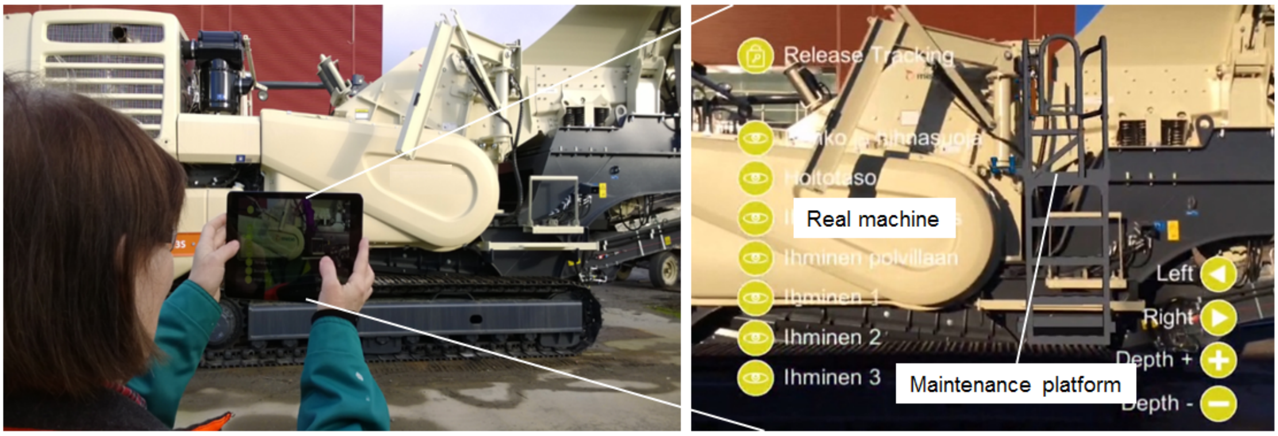

Figure 1. The augmented reality system for reviewing the maintenance platform. A participant is holding a tablet PC in her hand on the left-hand side. On the right-hand side is a screenshot from the tablet PC: a 3D model of a maintenance platform is augmented on top of the real machine.

In the VE test system, the model included a virtual model of the product (the maintenance platform), a virtual model of the rock crushing machine, a virtual environment, three different postures of DHM, a real participant and 3D models of hands and shoes (Figure2). The models of the maintenance platform, the machine and DHMs did not have dynamic model characteristics. A participant's head, hands and feet were tracked, and therefore the participant was able to move hand and shoe models in the VE. The participant was able to stand next to the rock crushing machine or on top of the maintenance platform. He/she was able to walk around. A user interface consists of the means of manipulation and sensory modalities. The means of manipulation were provided to the participant by using verbal commands to show and hide DHMs, and to change the standing location (the Wizard of Oz approach). The participant was also able to move around. Only visual cues were provided as a sensory modality. Therefore, haptic feedback was not provided. Nevertheless, the participant was able to estimate collisions by using his/her hand and to look when it touched e.g. the railings of the maintenance platform. The hardware and software used were a head-mounted display (HMD) (Oculus), tracking (Vicon), Unity and Middle VR.

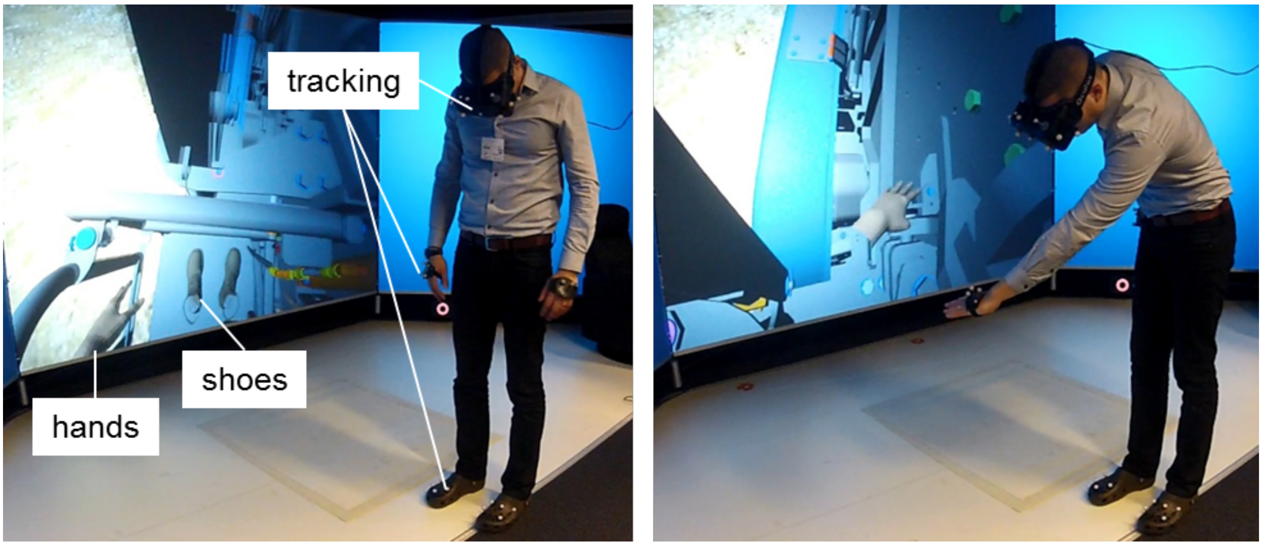

Figure 2. The virtual environment system for reviewing the maintenance platform. Other people were able to see where the participant is looking from the screen behind.

2.5 Test procedure

The AR study was conducted at the company facilities and the VE study in a VR-laboratory. At the beginning, there was an introduction to the project and the test to participants. The participants' goal was to review the possibility of performing two maintenance tasks on the maintenance platform: (1) a visual check of a feeder of the rock crushing machine, and (2) attempt to open a bolt in the machine frame. The use case was an upgrade model of a maintenance platform for a mobile rock crushing machine. Next, a consent form and the participant demographics were collected. After gathering initial information, the participant went to the test area (outside or VR lab) and was given a short introduction to the use of the AR system or the VE system. The participant was able to try the system for a while. In the AR system, the participant placed the maintenance platform in the right place, checked if it was possible to perform the maintenance tasks mentioned in the introduction and ended the task. In the VE system, the model of the maintenance platform was already in the correct place. After reviewing the maintenance task, the participant was asked to complete Nasa-TLX with unweighted scores, and other questionnaires about the overall assessment of a design object and the suitability of the system used for HFE evaluation. The interviewer wrote down the answers to the open questions and explanations of the selected Likert scale results. In addition, a modified SSQ was collected before and after the test with the VE system. The whole test took around 45 minutes in total including data collection.

3. Results

T-test was used to analyse differences between the two virtual prototypes in: overall assessment of the maintenance platform, suitability for the HFE evaluation, and subjective experience of workload and simulator sickness. Results indicate that both AR and VE prototypes were suitable to support HFE evaluation. However, the VE system was valued as being more suitable to support visibility and reach evaluation, and the assessment of the use of tools.

3.1 Overall assessment of the maintenance platform with the virtual prototypes

Seven people (n=9) said that they did not find any design issues from the maintenance platform while they used the AR system e.g. "I didn't find any design issues because I wasn't able to see things I would have wanted" (male design engineer, 47 years of age). Five test subjects said that they did not come up with any new ideas of how to improve the maintenance platform. Some other comments were also made such as a participant who would have wanted more information about the maintenance platform e.g. material choices, measurements (e.g. a railing height), attachments and dynamics (e.g. how the port opens). The participants felt that many things which are needed in design decision making were left uncertain, e.g. the correct placement of the maintenance platform, how the maintenance worker fits into the platform, is he/she able to reach the targets, and how other parts move near the platform.

The use of a VE system provoked more comments about the design issues than the AR system. Three (n=10) people said that they did not find any design issues with the maintenance platform. Four people mentioned that it was easy to see the feeder in the machine. Four people said that they were able to reach the bolt and perform the task. However, seven people said that the maintenance platform was tight and small: "The platform is small for big man" (male design engineer, 47 years of age). Four also mentioned that the reach distances in some cases are too long, e.g. when putting tools on the maintenance platform from the ground. When asked how to develop the maintenance platform, five people commented that the maintenance platform is OK. Nevertheless, five people commented that the maintenance platform could be bigger, and three people that the platform could be located a little lower. Four people would have wanted more information about design constraints: "I'm not able to say how to develop the maintenance platform further because I don't know all the constraints that are affecting the design" (male design engineer, 32 years of age). Three people did not require more information. Some other single comments were made about moving e.g. climbing stairs, going through the gate, kneeling on the platform.

With both prototypes, participants agreed that the maintenance platform is good, safe and efficient to work on the Likert scale (Table 1).

Table 1. The overall assessment of the maintenance platform, ($p<0.05$).

	AR (n=9)	VE (n=10)		
	M (SD)	M (SD)	t	p
The maintenance platform is good	3.78 (0.67)	3.70 (0.95)	0.204	0.840
The maintenance platform is safe to use	4.33 (0.71)	4.50 (0.71)	0.513	0.615
It is efficient to work on the maintenance platform	3.67 (0.71)	3.90 (0.74)	0.702	0.492

3.2 Suitability of the virtual prototypes for the human factors/ergonomics evaluation

The suitability of the virtual prototypes for the HFE evaluation was analysed with the questions from the HFE checklists as in Karwowski (2006). Totally, eleven questions were asked related to the maintenance task performed on the maintenance platform. In data analysis, two questions regarding visibility were put together as one to conclude a one HFE feature. The same was done to the two questions regarding reach issues.

Both systems received more than the mean value on the Likert scale in visibility, climbing, enough room, postures, use of tools and reach (Table 3). In addition, the VE system received the mean value in task performing time. Below the mean were environmental factors, force and task time (in the AR system). There was a significant difference in the visibility scores for the AR system (M= 3.3, SD=1.3) and the VE system (M=4.8, SD=0.4) conditions; t(22)=-5.66, p=0.000. There was significant difference in the reach scores for the AR system (M=3.39, SD=0.78) and the VE system (M=4.45, SD=0.61) conditions; t(36)=-4.721, p=0.000. In addition, there was significant difference in the use and carry tools for AR system (M=3.22, SD=0.83) and VE system (M=4.20, SD=0.79) conditions; t(17)=-2.62, p=0.018. Postures (p=0.059) and climbing (p=0.068) were not statistically different but the values were close.

Table 3.The suitability of the AR and VE systems for the human factors / ergonomics evaluation, (p<0.05).

By using this system I was able to evaluate…	AR (n=9)	VE (n=10)		
	M (SD)	M (SD)	t	p
the visibility of the feeder, the crusher and the frame bolt.	3.33 (1.03)	4.80 (0.41)	5.656	**0.000**
the load of the working postures.	3.11 (0.33)	3.7 (0.82)	2.080	0.059
the room for different working postures.	3.56 (1.01)	4.00 (1.05)	0.934	0.363
the reaches to the frame bolt from above and between railings.	3.39 (0.78)	4.45 (0.61)	4.721	**0.000**
the ergonomics and safety when climbing stairs.	4.00 (0.71)	3.20 (1.03)	1.947	0.068
the force needed to open the frame bolt.	2.00 (0.71)	2.10 (1.29)	0.206	0.839
the use of tools, and carrying them.	3.22 (0.83)	4.20 (0.79)	2.627	**0.018**
environmental effects (e.g. dust, noise, temperature).	2.67 (0.87)	2.20 (0.92)	1.136	0.272
the time spent opening the frame bolt.	2.78 (0.67)	3.00 (1.33)	0.451	0.658

3.3 Subjective experience of workload and simulator sickness

There was no significant difference between the AR and VE systems in subjective workload (Figure 3). Both systems were below mean value on the NASA-TLX scale (0 low demand, 100 high demand). In general, VE system received higher scores on mental, physical and temporal demand, but lower scores on performance, effort and frustration compared to AR system. None of these results were statistically significant. However, performance was close (p=0.054).

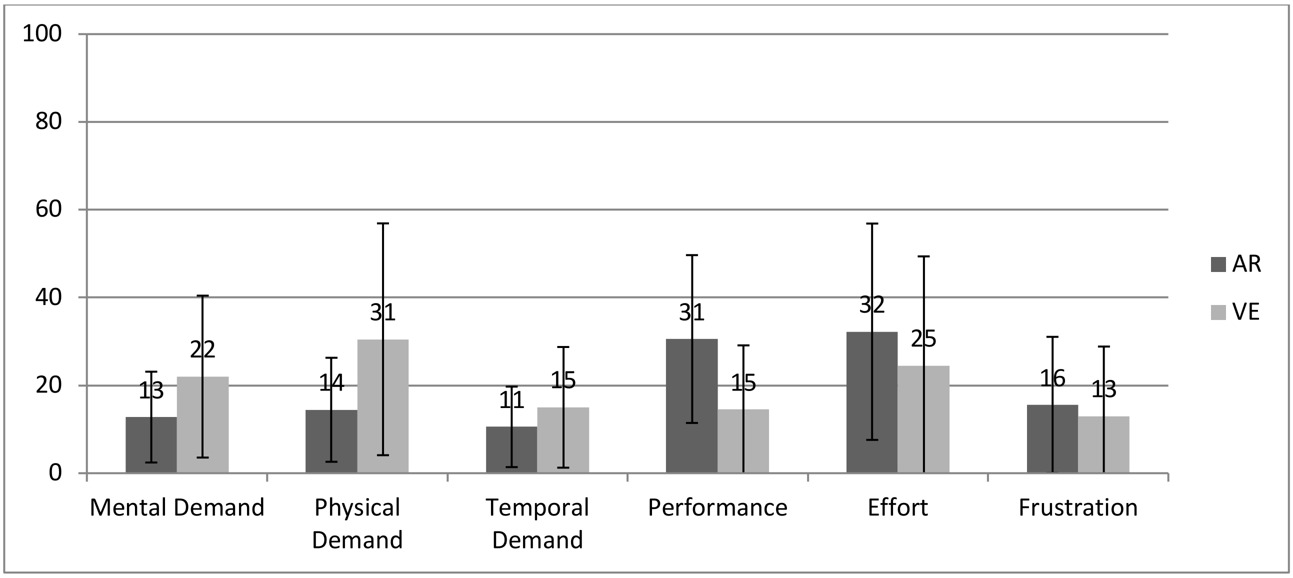

Figure 3. Subjective workload after using the AR and VE systems.

An adopted SSQ was used in the VE test to find out if negative symptoms appear. If the symptoms are moderate or severe, they can affect the performance and experience of the participant. However, there were no severe symptoms after being in VE (mean values between 1.00–1.44). The biggest change was in general discomfort but this was not significant either (before M=1.11, after M=1.44). The AR system was not so immersive, and therefore SSQ was not used in the AR test.

4. Discussion

Results indicate that VP can be used to support HFE evaluation during the design. The significant differences between the use of AR and VE systems for the HFE evaluation are related to the system model characteristics (fidelity, virtuality, dynamics, statics, etc.), and how the means of manipulation and sensory modalities (haptics, aural, etc.) are provided.

The system model characteristics impact the suitability of the virtual prototypes for the HFE evaluation. A significant difference in visibility would have been even more different if the DHMs had not been provided in the AR system. By comparing the DHM's head location to the machine's frame, participants were able to estimate whether a maintenance worker is able to see targets. In addition, the significant difference would have decreased by providing more information e.g. DHMs' field of views to the tablet PC in the AR system. This same analogy can be found from the results of the posture evaluation. It is remarkable that there was no significant difference between the AR and VE system when evaluating the postures. This is because DHMs with three static postures were provided in the AR system to support participants' HFE evaluation. In addition, the use of DHMs increased the reach evaluation results (M=3.39) in the AR system. However, reach was significantly different between the AR and VE systems. These findings suggest that it is important to provide required design information in some form to support design decisions making.

A natural and interactive interface with the context in use supports the HFE evaluation in VP. Based on the findings, the VE system was more natural and interactive than the AR system. A significant difference in visibility results was due to the fact that the participants were able to stand on the maintenance platform and see the feeder and other parts properly. In addition, in the VE system participants were able to use their own hands to see how far they can reach and also visually check whether their hands cut the model parts. Therefore, it was significantly better to do the evaluation of reach distances in the VE system. However, designers should remember that they are comparing the system dimensions against their own body size and this is not covering a whole population. The VE system also supported the evaluation of carrying and using tools (a significant difference from the AR system). In the VE, participants were able to perform tasks and act in the context. They were able to reach the maintenance platform from the ground and put the tools on top of it even though the tools did not exist virtually or physically. They were also able to imagine the use of tools while opening the bolt in the maintenance task. In the AR system, the interaction with the maintenance platform was limited because participants needed to keep the tablet PC in their hands and only a visual feedback was provided. In addition, the AR-technology was not robust enough to support more free movements. The image in the tablet PC was also small, so when you stood next to the rock crushing machine you were able to see only a small part of the maintenance platform. Therefore, participants tended to look at the maintenance platform from further away from the machine.

Other sensory modalities, in addition to the visual feedback, are required to better evaluate environmental factors, force and task time. The participants felt that the virtual prototypes were less suitable for the evaluation of environmental factors (e.g. noise, lighting) (AR M=2.7; VE M=2.2), force (AR M=2.0; VE M=2.1), and task time (AR M=2.8; VE M=3.0), from all HFE factors listed. Two different environments were provided: an outdoor environment and a virtual environment. However, none of these were real rough mine environments with noise and dust. Haptic and aural feedback was not provided in both cases and in the AR system the natural interaction with the maintenance platform was limited. However, sometimes the use of haptics does not increase the feeling of presence. Pontonnier et al. (2013) discovered that the mechanical limitations of the haptic device lowered the sensation of presence and an increase in the difficulty was reported compared to real environment and VEs. The results of the evaluation of the time required for the task could have been influenced by the fact that the participants were not required to perform the task step-by-step. Therefore, it might have been difficult to evaluate the time spent. According to Sauer and Sonderegger (2009), the task completion time may be overestimated when a computer- based prototype is used as compared to paper prototype.

Based on the subjective workload evaluation result, it can be seen that both systems were very usable and workload was not high. There was no significant difference between the systems. However, it can be seen that the use of a VE system was mentally, physically and temporally more demanding. Six participants (n=10) had not used the VEs before and this can have affected the results. Hu et al. (2011) compared VE and physical prototypes and also discovered that more effort is needed to perform a task in a VE than in a real environment. In addition, as a result of using HMD, participants felt more immersed and therefore experienced some SSQ symptoms. On the other hand, the use of the AR system required more effort, performance and was more frustrating. One reason for this could be the technology readiness and usability of the AR system. In addition, the experience was different within the AR system; it required more imagination and lacked natural interaction. This could have generated more frustration.

The different prototypes did not affect the overall assessment (good, safe and efficient to work) of the maintenance platform. This supports the findings from the studies made about the system model characteristics of the prototype (not virtual prototypes, but e.g. computer vs. paper prototypes). In these studies, the fidelity level does not affect perceived usability, and therefore reduced fidelity prototypes are generally suitable to predict the product usability of the real appliance (Boothe et al., 2013)(Sauer and Sonderegger, 2009; Sauer et al., 2010). In addition, it also supports the Bruno and Muzzupappa's (2010) discover that VR techniques are valid alternative to traditional methods for product interface usability evaluation and that the interaction with the virtual interface does not invalidate the usability evaluation itself.

This research had limitations that may have affected to the validity of the results. The AR system had technology challenges: the 3D model in the tablet PC was not stable and it vibrated on the screen sometimes. In addition, it was not possible to freely walk around and view the maintenance platform model from different perspectives because the rotation in the AR system did not work correctly. Another validity issue may derive from the use of the between group setup and the difference in professions of the participants. However, all of the participants were stakeholders that usually could have participated in a design process at some point.

5. Conclusions

The purpose of this study was to evaluate the suitability of the VP to support HFE evaluation during the design. Two different prototype systems (AR and VE) were tested in a between-groups set up. Design engineers and other stakeholders evaluated the suitability of the prototypes to support the HFE evaluation.

Results indicate that both AR and VE prototypes were suitable for the assessment of visibility, climbing, postures, space, reach and use of tools. However, the VE system was valued as being more suitable to support the assessment of visibility, reach, and the use of tools than the AR system. To assess HFE factors such as environment, force and time, more sensory modalities, are required in addition to the visual feedback. Moreover, results show that the system model characteristics can impact the suitability of the virtual prototype for the HFE evaluation. A more natural and interactive interface with the context of use can support the HFE evaluation.

It is challenging to use complex systems such as virtual prototypes in the design process. The findings of this study can provide guidance for the use of virtual prototypes in HFE evaluation. In addition, when using virtual prototypes in design, it is not always important to go for high fidelity and high virtuality prototypes. It is more important to use virtual prototypes that provide enough details and information to make good design decisions.

Acknowledgements

The study was funded under the European Commission's Seventh Framework in the project Use-it-wisely (609027) "Innovative continuous upgrades of high investment product-services", and by VTT Technical Research Centre of Finland Ltd . The authors are grateful to all researchers and company representatives who have contributed to and supported the work presented in this publication. In addition, special thanks to D.Sc. (Tech.) Eija Kaasinen for her insightful comments on the paper and Iina Aaltonen for providing language help.

References

Aromaa, S., Leino, S., Viitaniemi, J., Jokinen, L., Kiviranta, S., 2012. Benefits of the use of virtual environments in design review meeting, in: Proceedings of the 12th International Design Conference. University of Zagreb, Dubrovnik, Croatia, p. 8 p.

Azuma, R., 1997. A survey of augmented reality. Presence Teleoperators Virtual Environ. 6, 355–385. doi:10.1.1.30.4999

Barbieri, L., Angilica, A., Bruno, F., Muzzupappa, M., 2013. Mixed prototyping with configurable physical archetype for usability evaluation of product interfaces. Comput. Ind. 64, 310–323. doi:10.1016/j.compind.2012.11.010

Boothe, C., Strawderman, L., Hosea, E., 2013. The effects of protoype medium on usability testing. Appl. Ergon. 44, 1033–1038. doi:10.1016/j.apergo.2013.04.014

Bordegoni, M., Caruso, G., 2012. Mixed reality distributed platform for collaborative design review of automotive interiors. Virtual Phys. Prototyp. 1–17. doi:10.1080/17452759.2012.721605

Bordegoni, M., Cugini, U., Caruso, G., Polistina, S., 2009. Mixed prototyping for product assessment: a reference framework. Int. J. Interact. Des. Manuf. 3, 177–187. doi:10.1007/s12008-009-0073-9

Bordegoni, M., Cugini, U., Ferrise, F., Graziosi, S., 2014. A method for bringing user experience upstream to design. Virtual Phys. Prototyp. 00, 1–14. doi:10.1080/17452759.2014.934574

Broberg, O., Andersen, V., Seim, R., 2011. Participatory ergonomics in design processes: The role of boundary objects. Appl. Ergon. 42, 464–472. doi:10.1016/j.apergo.2010.09.006

Bruno, F., Muzzupappa, M., 2010. Product interface design: A participatory approach based on virtual reality. Int. J. Hum. Comput. Stud. 68, 254–269. doi:10.1016/j.ijhcs.2009.12.004

Bullinger, H.-J., Dangelmaier, M., 2003. Virtual prototyping and testing of in-vehicle interfaces. Ergonomics 46, 41–51. doi:10.1080/00140130303528

Canuto da Silva, G., Kaminski, P.C., 2015. Selection of virtual and physical prototypes in the product development process. Int. J. Adv. Manuf. Technol. doi:10.1007/s00170-015-7762-2

Cecil, J., Kanchanapiboon, A., 2007. Virtual engineering approaches in product and process design. Int. J. Adv. Manuf. Technol. 31, 846–856. doi:10.1007/s00170-005-0267-7

Chaffin, D., 2007. Human motion simulation for vehicle and workplace design. Hum. Factors Ergon. Manuf. 17, 475–484. doi:10.1002/hfm

Chaffin, D.B., 2005. Improving digital human modelling for proactive ergonomics in design. Ergonomics 48, 478–491. doi:10.1080/00140130400029191

Choi, S., Jung, K., Noh, S.D., 2015. Virtual reality applications in manufacturing industries: Past research, present findings, and future directions. Concurr. Eng. 23, 40–63. doi:10.1177/1063293X14568814

Demirel, H.O., Duffy, V.G., 2007. Application of digital human modeling in industry. Digit. Hum. Model. 824–832.

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W.S., Wilson, J.R., van der Doelen, B., 2012. A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics 55, 377–95. doi:10.1080/00140139.2012.661087

Eastgate, R.M., Wilson, J.R., D'Cruz, M., 2014. Structured Development of Virtual Environments, in: Hale, K.S., Stanney, K.M. (Eds.), Handbook of Virtual Environments: Design, Implementation, and Applications. CRC Press, pp. 353 – 389.

Ferrise, F., Bordegoni, M., Cugini, U., 2013. Interactive virtual prototypes for testing the interaction with new products. Comput. Aided. Des. Appl. 10, 515–525. doi:10.3722/cadaps.2013.515-525

Feyen, R., Liu, Y., Chaffin, D., Jimmerson, G., Joseph, B., 2000. Computer-aided ergonomics: A case study of incorporating ergonomics analyses into workplace design. Appl. Ergon. 31, 291–300. doi:10.1016/S0003-6870(99)00053-8

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., Tecchia, F., 2013. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interact. Learn. Environ. 4820, 1–21. doi:10.1080/10494820.2013.815221

Hall-Andersen, L.B., Broberg, O., 2014. Integrating ergonomics into engineering design: The role of objects. Appl. Ergon. 45, 647–654. doi:10.1016/j.apergo.2013.09.002

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv. Psychol. 52, 139–183.

Hu, B., Ma, L., Zhang, W., Salvendy, G., Chablat, D., Bennis, F., 2011. Predicting real-world ergonomic measurements by simulation in a virtual environment. Int. J. Ind. Ergon. 41, 64–71. doi:10.1016/j.ergon.2010.10.001

Huet, G., Culley, S.J., McMahon, C.A., Fortin, C., 2007. Making sense of engineering design review activities. Ai Edam-Artificial Intell. Eng. Des. Anal. Manuf. 21, 243–266. doi:10.1017/s0890060407000261

IEA, 2000. Definition and domains of ergonomics [WWW Document]. URL http://www.iea.cc/whats/ (accessed 1.15.16).

ISO 9241-210, 2010. Human-centred design for interactive system. Geneva.

Jia, D., Bhatti, A., Nahavandi, S., Horan, B., 2012. Human Performance Measures for Interactive Haptic-Audio-Visual Interfaces. IEEE Trans. Haptics 1–1.

Kalawsky, R.S., 1993. The Science of Virtual Reality and Virtual Environments. Addison-Wesley.

Karkee, M., Steward, B.L., Kelkar, A.G., Kemp, Z.T., 2011. Modeling and real-time simulation architectures for virtual prototyping of off-road vehicles. Virtual Real. 15, 83–96. doi:10.1007/s10055-009-0150-1

Karwowski, W., 2006. The Discipline of Ergonomics and Human Factors, in: Salvendy, G. (Ed.), Handbook of Human Factors & Ergonomics. John Wiley, New York, pp. 1–25.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. 3, 203–220.

Kremer, K., 1998. A concept for virtual reality tools for design reviews, in: Proceedings of the Conference on Visualization '98. IEEE, Research Triangle Park, NC, USA, pp. 205–210. doi:10.1109/VISUAL.1998.745304

Kulkarni, a, Kapoor, a, Iyer, M., Kosse, V., 2011. Virtual prototyping used as validation tool in automotive design. Built Environ. 12–16.

Lawson, G., Herriotts, P., Malcolm, L., Gabrecht, K., Hermawati, S., 2015. The use of virtual reality and physical tools in the development and validation of ease of entry and exit in passenger vehicles. Appl. Ergon. 48, 240–251. doi:10.1016/j.apergo.2014.12.007

Lawson, G., Salanitri, D., Waterfield, B., 2016. Future directions for the development of virtual reality within an automotive manufacturer. Appl. Ergon. 53, 323–330. doi:10.1016/j.apergo.2015.06.024

Leino, S.-P., 2015. Reframing the value of virtual prototyping. Intermediary virtual prototyping - the evolving approach of virtual environments based virtual prototyping in the context of new product development and low volume production. VTT, Espoo.

Lim, Y.-K., Pangam, A., Periyasami, S., Aneja, S., 2006. Comparative analysis of high- and low-fidelity prototypes for more valid usability evaluations of mobile devices. Proc. 4th Nord. Conf. Human-computer Interact. Chang. roles - Nord. '06 291–300. doi:10.1145/1182475.1182506

Lämkull, D., Hanson, L., Örtengren, R., 2009. A comparative study of digital human modelling simulation results and their outcomes in reality: A case study within manual assembly of automobiles. Int. J. Ind. Ergon. 39, 428–441. doi:10.1016/j.ergon.2008.10.005

Ma, D., Zhen, X., Hu, Y., Wu, D., Fan, X., Zhu, H., 2011. Simulation and Its Application, in: Virtual Reality & Augmented Reality in Industry. Springer, Berlin Heidelberg, pp. 55–82.

Mahdjoub, M., Al Khatib, A., Bluntzer, J., Sagot, J.-C., 2013. Multidisciplinary Convergence About " Product - Use " Couple: Intermediary Object 's Structure, in: ICED13: 19th International Conference on Engineering Design. Seoul.

Milgram, P., Takemura, H., Utsumi, A., Kishino, F., 1995. Augmented reality: A class of displays on the reality-virtuality continuum, in: Photonics for Industrial Applications. International Society for Optics and Photonics, pp. 282–292.

Mujber, T.S., Szecsi, T., Hashmi, M.S.J., 2004. Virtual reality applications in manufacturing process simulation. J. Mater. Process. Technol. 155-156, 1834–1838. doi:10.1016/j.jmatprotec.2004.04.401

Muller, M.J., Kuhn, S., 1993. Participatory design. Commun. ACM 36, 24–28.

Nee, A.Y.C., Ong, S.K., Chryssolouris, G., Mourtzis, D., 2012. Augmented reality applications in design and manufacturing. CIRP Ann. - Manuf. Technol. 61, 657–679. doi:10.1016/j.cirp.2012.05.010

Park, H., Moon, H.C., Lee, J.Y., 2009. Tangible augmented prototyping of digital handheld products. Comput. Ind. 60, 114–125. doi:10.1016/j.compind.2008.09.001

Perez, J., Neumann, W.P., 2015. Ergonomicsts' and engineers' views on the utility of virtual huma factors tools. Hum. Factors Ergon. Manuf. Serv. Ind. 25, 279–293.

Pontonnier, C., Dumont, G., Samani, A., Madeleine, P., Badawi, M., 2013. Designing and evaluating a workstation in real and virtual environment: toward virtual reality based ergonomic design sessions. J. Multimodal User Interfaces 1–10. doi:10.1007/s12193-013-0138-8

Sanjog, J., Karmakar, S., Patel, T., Chowdhury, A., 2015. Towards virtual ergonomics: aviation and aerospace. Aircr. Eng. Aerosp. Technol. 87, 266–273. doi:10.1108/AEAT-05-2013-0094

Sauer, J., Seibel, K., Rüttinger, B., 2010. The influence of user expertise and prototype fidelity in usability tests. Appl. Ergon. 41, 130–140. doi:10.1016/j.apergo.2009.06.003

Sauer, J., Sonderegger, A., 2009. The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. Appl. Ergon. 40, 670–677. doi:10.1016/j.apergo.2008.06.006

Seth, A., Vance, J.M., Oliver, J.H., 2011. Virtual reality for assembly methods prototyping: A review. Virtual Real. 15, 5–20. doi:10.1007/s10055-009-0153-y

Shen, Y., Ong, S.K., Nee, A.Y.C., 2010. Augmented reality for collaborative product design and development. Des. Stud. 31, 118–145. doi:10.1016/j.destud.2009.11.001

Sutcliffe, A., Gault, B., 2004. Heuristic evaluation of virtual reality applications. Interact. Comput. 16, 831–849. doi:10.1016/j.intcom.2004.05.001

Wang, G.G., 2002. Definition and Review of Virtual Prototyping. J. Comput. Inf. Sci. Eng. 2, 232.

Wilson, J.R., D'Cruz, M., 2006. Virtual and interactive environments for work of the future. Int. J. Hum. Comput. Stud. 64, 158–169. doi:10.1016/j.ijhcs.2005.08.007

Wu, T., Tian, R., Duffy, V.G., 2012. Performing Ergonomics Analyses through Virtual Interactive Design: Validity and Reliability Assessment 22, 256–268. doi:10.1002/hfm