IDENTIFYING THE EFFECT OF REVIEWERS' EXPERTISE ON DESIGN REVIEW USING VIRTUAL REALITY AND DESKTOP INTERFACE

N. Horvat ✉, S. Škec, T. Martinec, F. Lukačević and M. M. Perišić

University of Zagreb, Croatia

✉ horvatn@fsb.hr

Abstract

This paper suggests that analysing the effect of visualisation technologies during design reviews should include variables related to design reviewers' expertise and focus on the process variables rather than the outcomes. The experiment showed better averages in terms of design understanding for groups in desktop interface than for groups in virtual reality. However, the observed difference might also be due to experience with the technology. Finally, regardless of the visualisation technology, higher expertise group showed consistently better design understanding than lower expertise groups.

Keywords: virtual reality (VR), design review, visualisation, computer-aided design (CAD), design expertise

1. Introduction

Design reviews are one of the product development's control mechanisms which serve to validate existing solutions, verify the quality of conducted work, manage upcoming activities and propose new solutions (D'Astous et al., 2004; Huet et al., 2007). Researchers emphasise the importance of design reviewers' (hereinafter reviewers) expertise (Huet et al., 2007; Liu et al., 2018; Thimmaiah et al., 2016) and shared design understanding (D'Astous et al., 2004; Huet et al., 2007; Liu et al., 2018) for the effective design review. To aid in sharing the understanding of the design, reviewers use different types of design representation (e.g. 3D models, drawings, documents). Higher fidelity of these representations results in increased confidence and effectiveness during design reviews (Hannah et al., 2012). A common type of high fidelity representation is a 3D model, which reviewers use as the main artefacts of design review (Di Marco et al., 2012). Given the importance of 3D models and fidelity of design representation, it is not surprising that researchers often study the use of virtual reality (VR) as a support tool for design reviews (Coburn et al., 2017).

The idea behind VR is to artificially stimulate human sensors with little or no awareness of the interface. For example, VR with position tracking and stereoscopic display enables at least two depth cues – motion parallax and stereo depth cue – which are typically not stimulated by other visualisation technologies such as desktop interface (DI). Furthermore, this technology facilitates 3D manipulation of the objects, thus enabling the use of proprioceptive sense while reviewing the model. Accordingly, researchers concerned with the application of VR in design reviews identified better spatial perception while comparing VR technologies with DI (Horvat et al., 2019; Paes et al., 2017). On the contrary, studies investigating the effect of visualisation technology on design review outcomes (e.g. the number of identified errors) showed inconclusive findings (de Casenave and Lugo, 2017, 2018). We

argue that these studies oversimplify the number of studied variables and that the relationships between input and output variables might be visible through design review process variables (e.g. reviewers' behaviour or cognitive state). The research on the effect of visualisation technologies on design reviews is still in its early phases. Therefore, the identification of influential variables is necessary for the assessment of appropriate visualisation technology in design reviews. For example, previous research proposed that experience might aid spatial perception in design reviews.

Given the importance of design understanding, and the relevance of reviewers' expertise in design reviews, we have studied the effect of expertise on design understanding using VR and DI. The following research questions guided the study of the aforementioned effect:

- Does the visualisation technology affect design understanding?
- Does the reviewers' expertise affect design understanding?

The paper is organised in six sections. After the introduction section, we have presented the research background on design reviews, visualisation of 3D models and expertise in Section 2. We have then described the conducted empirical study in Section 3, presented the results in Section 4 and discussed them in Section 5. Finally, we have given conclusions of the study in Section 6.

2. Background

2.1. Design review process

When preparing a design review meeting, a responsible person collects and sends the necessary information to the participants (Sater-Black and Iversen, 1994). Having documents in advance helps reviewers develop their understanding of design. This importance of understanding is captured in descriptive studies, as they suggest that shared understanding is a prerequisite for an evaluation (D'Astous et al., 2004). Therefore, sharing information and understanding design intent are essential activities in design reviews (Huet et al., 2007; Liu et al., 2018). Indeed, empirical studies revealed that the amount of shared information affects the ability to identify design problems (Wetmore et al., 2010). Furthermore, empirical studies proposed the use of high-fidelity representations in design reviews, as they increase confidence and correctness in reviewers' conclusions (Hannah et al., 2012).

Building on the importance of sharing information and understanding design, researchers proposed tools which enable the use of VR for reviewing 3D models. For example, Freeman et al. (2018) developed an interface between commercial Computer-aided design (CAD) package and VR environment with features such as hide, exploded view, a parametric update of the model and colour model. Similarly, Wolfartsberger (2019) proposed a VR-based tool which reads Filmbox (*.fbx*) files and has several interactions mechanisms: Looking and walking, teleporting, touching and highlighting, grabbing objects, separating and merging objects, and hiding and showing objects. Following the development of supporting tools for VR-based design reviews, another stream of research focused on understanding the differences between commonly used visualisation technologies such as DI and VR.

2.2. The effect of VR on understanding 3D models

Studies which examined the effect of visualisation technology on design reviews mainly focused their observations on the review outcomes (e.g. the number of identified design errors (Wolfartsberger, 2019)). However, these studies reported contradictory findings, e.g. some studies found an increased number of identified design errors in a VR environment, while others found a decreased number of the identified errors. As a response, rather than focusing on design review outcomes, recent studies investigated the effect of visualisation tools on various design review process variables.

Self-reported studies found that participants see the benefit of using immersive technologies for understanding the spatial relationship between parts (Berg and Vance, 2016). Similarly, experiments which investigated the effect of visualisation found consistently better spatial perception when using VR. For example, some of the reviewed studies identified significant improvement in spatial perception when participants (either students or professionals) used stereoscopic environments as compared to non-stereoscopic ones (Horvat et al., 2019; Paes et al., 2017; Satter and Butler, 2015). Furthermore, a recent

study suggested that product complexity levels might influence spatial perception, as differences between VR and DI were more noticeable at higher complexity levels (Horvat et al., 2019).

The use of different visualisation technologies (VR and DI) showed consistent differences in terms of spatial perception. It is hence a logical step forward to identify if the use of different technologies affects the understanding of a design artefact. Researchers usually measured design understanding by investigating the correctness of the conducted task. In this stream, two VR toolsets with different functions have exhibited different results in terms of participants' correctness and confidence in geometry understanding (Freeman et al., 2018). However, in a comparison of VR and DI visualisation, de Casenave and Lugo (2017) did not find significant differences in task scores in-between the environments. They even found that participants spent more time in a VR-based review since they had more interaction mechanisms than in DI-based review.

It is still not clear how current technologies affect design understanding as only a few studies have investigated the topic. Moreover, findings of these studies are not as consistent as findings of the studies which investigated the effect of technology on spatial perception, suggesting that comparative studies did not control for some influential variables. A recent study indicated that experience might affect spatial perception (Horvat et al., 2019), suggesting that the reviewers' expertise could give more insights in empirical studies.

2.3. Design reviewers' expertise

Design reviews serve as design control mechanisms (Huet et al., 2007) and result in suggestions related to design artefact and design process. Therefore, to a certain extent, a six-level model of designers' expertise (Lawson and Dorst, 2009, p. 99) can also describe a reviewers' expertise. The model consists of a novice, advanced beginner, competent, expert, master and visionary designer (Lawson and Dorst, 2009, p. 99). Although design researchers usually identify students as a novice, Lawson and Dorst (2009, p. 103) reported that some of the students were even at the competent level. Therefore, even in the studies with students, the expertise might have important implications on the observed phenomena. To acquire expertise, one of the key factors is deliberate practice and the accumulation of experience (Cross, 2004). Therefore, it is common to assess designers' expertise by assessing their years of experience.

However, besides design related expertise, it is often necessary to also have contextual expertise (Lawson and Dorst, 2009, p. 84). In terms of contextual expertise, previous studies found that participants who had high user and product familiarity did not have problems while generating requirements (Morkos and Summers, 2013). Moreover, these participants also generated requirements at a faster rate than participants with low contextual experience (Morkos and Summers, 2013). In an ideation activity, Hu and Reid (2018) found a negative effect of contextual experience on mental states associated with creativity. They also studied the effect of contextual experience on the quantity and quality of design solutions, but they did not find any effect (Hu and Reid, 2018). Therefore, to get a better overview of designers' expertise, researchers often use this combination of design domain experience and contextual experience (consult Mosely et al. (2018) for an example).

Since researchers found that VR has a consistent effect on spatial perception (Horvat et al., 2019; Paes et al., 2017; Satter and Butler, 2015), spatial thinking might be a relevant individual ability for performing design reviews. This ability can be divided into four categories (Newcombe and Shipley, 2015) distinguished by intrinsic (within an object) or extrinsic (between object), and static (coding of spatial information) or dynamic (transforming the spatial information). Commonly studied category is intrinsic-dynamic with the use of mental rotations (e.g. the Mental Rotation Test (MRT) (Peters et al., 1995)). However, previous research failed to identify its effect on the level of design understanding (de Casenave and Lugo, 2017).

Previous studies showed that VR has great potential to support the design review process. Nevertheless, inconsistent findings of the effect of visualisation technologies have a consequence of still undefined VR application in design reviews, i.e. for which products, in what situation and in what manner the VR technology should be used. In addition, although the effect of expertise in design is widely acknowledged, studies of VR-based design reviews have often ignored the expertise of the reviewers (especially when the studied participants are solely students). Therefore, analysing the effect

of visualisation technologies in conjunction with reviewers' expertise on design understanding might reveal new findings towards assessing the use of VR in the design field.

3. Empirical study description

The empirical study aims to identify the effect of visualisation technologies and reviewers' expertise on design understanding. The study follows a quasi-experimental design with independent samples between VR and DI conditions. We operationalised reviewers' expertise through the intrinsic-dynamic type of spatial thinking and experiences related to technology, design domain, and context. The intrinsic-dynamic type of spatial reasoning was measured with a Mental Rotations Test (MRT) (Peters et al., 1995). The level of this reasoning was used to allocate participants in one of the two groups ensuring a similar distribution of spatial reasoning abilities in both environments. We measured experiences using a prior experience questionnaire (PEQ) with 24 items. Table 1 shows the mapping of the PEQ items with variables studied in this paper. Most of the variables relate to only one item from the PEQ. The exception is a DR technology experience variable, which consists of VR experience item for the VR group, and CAD experience item for DI group. To measure differences between the two groups, we developed eight tasks, each related to a different aspect of design review, but herein analysed only one of them.

Table 1. Variables from prior experience questionnaire (PEQ)

Category	Example questions	Scale	Variable name
Technology	-How much experience do you have in using CAD?	None (1) – ≥5 (5)	DR technology experience
	- How much time have you spent in Virtual Reality?	Interval [min]	
Design	-How many design reviews have you conducted?	None (1) – ≥10 (5)	Conducted DRs
	-How many years of experience do you have in mechanical design?	None (1) – ≥6 (5)	ME experience
	- How many mechanical design projects, including student projects, have you worked on?	None (1) – ≥10 (5)	ME projects
Contextual	-How many times have you folded a one-person foldable vehicle?	None (1) – >30 (5)	Number of vehicle foldings
	-How many times have you adjusted height or angle of a seat on bicycle, tricycle or scooter?	None (1) – >30 (5)	Number of seat adjustments
	-How many times have you adjusted height or angle of a steering bar on bicycle, tricycle or scooter?	None (1) – >30 (5)	# of steering bar adjustments

3.1. Study participants and materials

Forty (40) students from two universities (24 students at one university and 16 students at another) participated in the study. At each university, half of the participants conducted the task in VR, while the other half in DI. In total, 33 male students and seven female students participated in the study, ranging in age from 19 to 31 (22.25 ± 2.47) in DI and 20 to 27 (22.6 ± 1.91) in VR. The use of students might limit the generalisation of findings. However, students can develop the expertise of competent designers (Lawson and Dorst, 2009, p. 103) and studying them might provide some practical implications.

The product under review was a foldable mobility scooter, a representative of the products at the third complexity level (Hubka and Eder, 1988). The scooter was developed to the state of a functional prototype in a company-university collaboration. The scooter is supposed to fold into the size that can fit in a car boot and resemble a rolling suitcase, to be as easy to manoeuvre as possible and to be light so a single person can carry it. During the experiment, participants had available two states of the scooter: driving mode and folded mode.

DI setup for the design review session included a high-performance computer (e.g. Intel Xeon 4114 CPU, 32 GB DDR4, and NVIDIA Quadro P5000 16GB GDDR5X), a 24'' monitor (resolution 1920x1080), and conventional input devices (mouse and keyboard). A CAD software to present 3D models was Siemens NX, versions 10 and 12. VR setup included the same computer with an HTC Vive PRO headset and Autodesk VRED PRO 2020. The Autodesk software included scripts for interaction such as a visualisation of a 30 cm ruler, section tool, rotation of the model, switch between the driving and folded modes and switch between scooter position (scooter on the ground and at the height of eyes; see Figure 1a).

Participants could also walk around the virtual room (size 1.9 m x 2 m) and thus navigate themselves. We recorded the audio of the participants' answers and the content displayed on the screen.
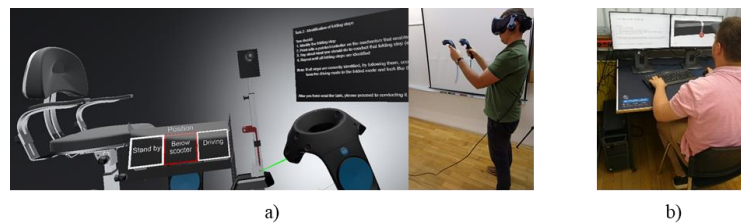

Figure 1. Identifying a folding step in: a) Virtual reality (VR), and b) Desktop interface (DI)

Before the experiment, we developed a presentation to guide the participants throughout the study and to control the amount of given information. The presentation included an introductory matter, a design review tutorial, and an introduction to tasks and questions. In the DI environment, four tutorial tasks and all other tasks were also presented on the left screen. In the VR environment, these tasks were inserted in the virtual environment as 3D models (see Figure 1a). Regardless of the presentation, a researcher was present during the experiment to help the participants with technical issues.

All the materials except the presence questionnaire, the technology tutorial and NX interface were bilingual (English and Croatian). One author translated the presentation, questionnaire, tasks and questions from English to Croatian, while another did a backward translation from Croatian to English.

3.2. Experimental procedure and analysed study task

At the beginning of the experiment, the researcher gave participants a brief introduction about the procedural steps. The researcher then gave an MRT to the participant, following the procedure provided with the test instructions. Next step encompassed a presentation-based design review tutorial. Then, the presentation guided the participants to complete a PEQ after which participants went through examples of tasks and questions - also on presentation. The last step before the design review session included the technology tutorial (based on the environment in which the participants conducted the design review). The design review session included one task at a time, which was related to different design review aspects (e.g. understanding, error detection). The procedure lasted for around 1.5h. To test the procedure, we conducted two pilot experiments, one with a participant being assigned to DI, and another one with participants working in VR.

The task of concern in this paper was the identification of folding steps to transform the scooter from driving to folded mode. We define a folding step as the relative motion of two subassemblies around one axis necessary to transform the scooter from one of the two states to another. We used the correctness of their identification as a proxy of understanding the design. In terms of the Function-Behaviour-Structure ontology (Gero and Kannengiesser, 2004), the participants had to identify behaviour from the structure, also defined as analysis. In total, six folding steps were necessary to fold the vehicle (Figure 2): Swivel rotation, Seat rotation, Backrest rotation, Seat holder rotation, Chassis rotation, and Tiller rotation. Participants had 3.5 minutes to finish the task.
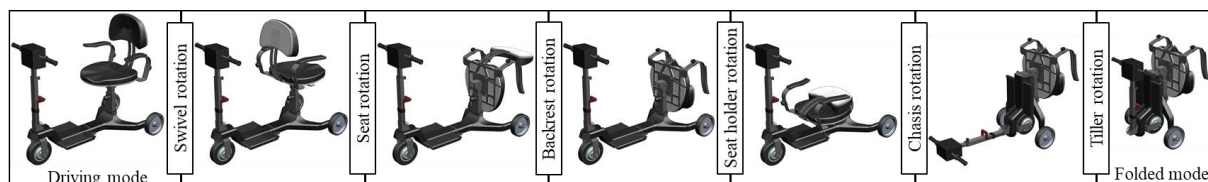

Figure 2. Steps in folding the one-person vehicle

4. Results

This section presents the results of the empirical study. All the participants identified at least one folding step, and only one participant identified all six folding steps (Figure 3). In DI, most of the participants identified five folding steps while in VR, most of the participants identified four folding steps. To get a

better insight into the differences between the two environments, we conducted inferential statistics by analysing each expertise variable (Table 1) in conjunction with the Environment variable (VR or DI). Firstly, we calculated the correlation between input variables to capture their interdependence (Section 4.1). Since the aim of this study is to identify variables which might affect the aspects of design review rather than determining their influence, we dichotomised the variables using a mean split. We are not stating that reviewers' expertise or any of its variables are binomial nor that we are capturing the extreme values of their expertise. We thus call the dichotomised groups simply Lower (L) and Higher (H), as opposed to commonly used low and high categorisation. To give more meaning into these groups, we usually presented intervals which fall in each of the groups. Section 4.2 presents the effect of studied variables on the number of correctly identified folding steps using a two-way ANOVA.
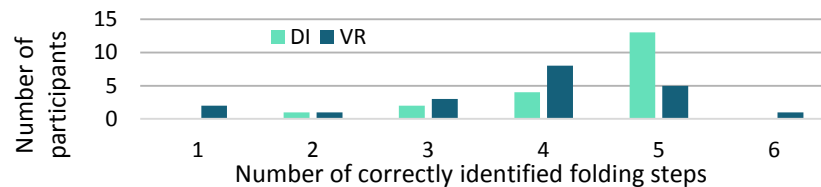

Figure 3. Distribution of correctly identified folding steps

4.1. Correlation between input variables

As most of the input variables are at the ordinal scale, Spearman's correlation served as a measure of association between input variables (Table 2). Results from the table suggest that many variables are linked, so the effect of each variable might be related to the effects of associated variables. For example, ME experience and ME projects variables correlate with all design experience variables and Number of seat adjustments. Additionally, ME projects variable also correlates with the Number of steering bar adjustments. The effect of ME variables might thus attribute to some of these variables or their combination. The Number of seat adjustments correlates with both ME variables, all contextual experience variables and with MRT score. MRT score correlates with Conducted DRs and Number of seat adjustments. Finally, Environment variable correlates with DR technology experience.

Table 2. Spearman's rank correlation coefficients between input variables

		Conducted DR	ME experience	ME projects	Number of vehicle foldings	Number of seat adjustments	Number of steering bar adjustments	MRT score	Environment
DR technology experience		0.322	0.256	0.212	0.046	0.254	0.167	0.201	-0.877
Design experience	Conducted DRs		0.318	0.268	0.059	0.260	0.264	0.296	-0.254
	ME experience			0.684	0.074	0.305	0.124	0.191	-0.035
	ME projects				-0.007	0.370	0.396	0.092	-0.042
Contextual experience	Number of vehicle foldings					0.333	0.243	0.014	0.000
	Number of seat adjustments						0.566	0.406	-0.014
	Number of steering bar adjustments							0.204	-0.087
Spatial thinking	MRT score								-0.105

DR – Design Review   ■ $p < 0.1$
ME – Mechanical Engineering   ■ $p < 0.05$
MRT – Mental Rotations Test   ■ $p < 0.01$

4.2. Correctly identified folding steps

To determine correctly identified folding steps, we compared participants' answers to the folding steps listed in Figure 2. Figure 4 provides the results as error bars (group mean and standard error) for each of the four categories (i.e. Lower and DI, Lower and VR, Higher and DI, and Higher and VR) within each combination of variables. In Figure 4, x-axis presents one variable of reviewers' expertise with the two levels for each metric (below the mean and above the mean) while the environment variable is colour coded. In addition to graph interpretation, a two-way ANOVA served as a statistical test to identify the significant variables. It is worth noting that all ANOVA models had a non-normal distribution of residuals, thus violating the assumption of normality. However, this violation has only a small effect on p-value (Glass et al., 1972).

The *p*-values presented on the top of each graph in Figure 4 show that the Environment variables have a significant effect within all the models. Therefore, the DI group had significantly higher means than their counterparts in VR. As shown in Figure 4, a group with higher DR technology experience had higher means within both environments. However, a two-way ANOVA showed that the effect of DR technology experience on the number of identified folding steps is not significant (*p*=0.13).
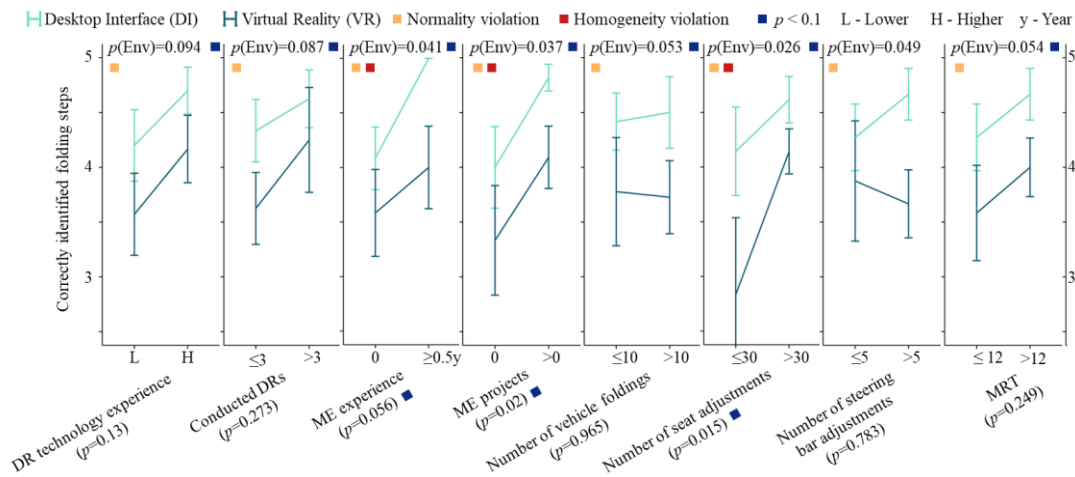

Figure 4. Error bars for correctly identified folding steps

All groups higher in design experience variables (i.e. Conducted DRs, ME experience and ME projects) also exhibited higher means in both environments. Two-way ANOVA showed that Conducted DRs do not significantly affect (*p*=0.273) mean scores of correctly identified folding steps. On the contrary, both ME experience and number of ME projects have a significant effect on the group means (p=0.056 and p=0.02, respectively). However, Levene test showed that ME experience and ME project variables violate the assumption of homogeneity. This violation is probably due to the ceiling effect as most of the participants identified five folding steps, while only one of them identified all six folding steps.

Furthermore, contextual expertise variables show a contradictory effect. Higher values in some variables (Number of seat adjustments) increased the mean of the response while in others response was environment-dependent (Number of vehicle foldings and Number of steering bar adjustments). Two-way ANOVA resulted in insignificant results for Number of vehicle foldings (*p*=0.965) and Number of steering bar adjustments (*p*=0.783). On the other hand, the Number of seat adjustments is statistically significant (*p*=0.015), but the Levene test again showed the violation of homogeneity. Finally, an MRT score suggests that its higher score results in higher mean for participants in both environments. However, a two-way ANOVA failed to identify significant effects (*p*=0.249).

5. Discussion

5.1. Difference between VR and DI

Environment variable highly correlates with the DR technology experience (see Table 2), meaning that any differences identified between the two studied environments can be due to the level of experience in the used technology. Therefore, it is not clear whether the improved identification of folding steps in DI (see Figure 4) is because of the CAD features or because of the higher experience of using this technology. Nevertheless, the findings of increased understanding in DI do not relate to previous results as de Casenave and Lugo (2017) did not find significant differences in terms of design understanding between VR and DI. Furthermore, the findings of correctly identified folding steps do not relate to consistently better spatial perception in VR (Horvat et al., 2019; Paes et al., 2017; Satter and Butler, 2015). This difference might be due to the differences in observed concepts, i.e. spatial perception and design understanding. In contrast to spatial perception, design understanding requires categorisation of sensory inputs (e.g. foldable vehicle design) into thought concepts (e.g. wheel, seat, handle) which are then compared to previous knowledge (Goldstone et al., 2018). While we argue that

spatial perception might help in design understanding as an aid in the categorisation of concepts, such a claim could not be supported by the results. Therefore, future studies should aim towards separating the two confounding variables to unravel the potential of VR technology for design understanding.

5.2. The effect of reviewers' expertise on design understanding

Expertise variables were interdependent (see Table 2), so the direct effect of each variable is not possible to identify. However, as an increase in most of the expertise variables increased the means of correctly identified folding steps (all except the two contextual variables), we suggest that expertise indeed affects design understanding. For example, while the DR technology experience did not have a significant effect, a *p*-value of its effect was relatively low. We thus suggest further investigation of the experience in the technology used for design reviews. Moreover, some variables (ME experience, ME projects, Number of seat adjustments) significantly affected group averages of correctly identified folding steps (see Figure 4). These results need further clarification because of the homogeneity violation. The violation might arise because of the difference in difficulties between the identification of five and six folding steps. Despite the many new questions raised, the identified significant effects are in accordance with the previous literature. More specifically, the ME experience and the ME projects are in line with Cross's (2004) suggestion that designers gather expertise through experience. Lawson and Dorst (2009, p. 92) also support the notion that learning new procedures and acquiring experience serve as shifters through levels of expertise. We thus propose that these two variables affect design understanding as the studied task is closely related to the ME field. Still, these two variables correlate with other variables (e.g. Number of seat adjustments) so it is not clear whether this combination with other variables or solely the ME variables affected design understanding.

The effect of contextual experience is inconclusive. On the one hand, higher experience in terms of the Number of seat adjustments resulted in a higher number of correctly identified folding steps, suggesting that designers benefit from the contextual experience. Previous studies for requirement elicitation task already showed the importance of contextual experience (Morkos and Summers, 2013) thus supporting this suggestion. However, as this variable correlates with other significant variables such as those related to ME, the results require further examination. On the other hand, higher groups in terms of the Number of vehicle foldings and the Number of steering bar adjustments had lower group average of correctly identified folding steps, but only in VR environment. These results contradict the theory of design expertise and can be interpreted in at least two ways. First, it might be that contextual experience does not affect the Number of identified folding steps and that a significant effect of the Number of seat adjustment is a result of some other interdependent variable or their combination (see Table 2). Contextual experience does not always enhance the design process, as identified by Hu and Reid (2018) in the ideation context. The second interpretation might be found in the lack of external validity of the studied variable since we developed the PEQ from theoretical assumptions without validation. It thus might be that the variables do not measure the relevant dimensions of the contextual experience.

In accordance with previously conducted studies which tried to relate spatial ability and design understanding (de Casenave and Lugo, 2017), this study did not reveal a significant effect of intrinsic-dynamic spatial thinking on the number of identified folding steps. These results contradict theoretical suggestions. Namely, spatial thinking deals with representing and mentally transforming spatial information, suggesting that they should be critical for design understanding (Newcombe and Shipley, 2015). However, it might be that in this case, its effect is not as high as the effect of other variables.

5.3. Limitations of the conducted study

Although the study enabled identification of several constituent variables of reviewers' expertise, there are limitations which might prevent generalisation. First, the study is conducted on students, thus having limited implications for the design practice. Although the results might generalise to some extent since students might reach competent designers level (Lawson and Dorst, 2009, p. 103), wider applicability of the results should be studied in the future. Reviewed design artefact might also limit the generalisation as reviewers' use different sensory cues while perceiving models of different levels of complexity. Next, as previous studies found significant differences in design understanding between different toolsets in VR (Freeman et al., 2018), the findings of this study are limited to a single toolset used in VR. Finally, part

of the procedure at one location and the complete procedure at the other was in English. Therefore, it might be that foreign language influenced participants' understanding of the task.

Related to data analysis, all ANOVA models with both significant variables violated the homogeneity assumption. As the violation of this assumption has a high effect on the Type I error (Glass et al., 1972), the obtained p-values need further clarification. Finally, this study analysed several ANOVA models for various input variables, which increased the probability of Type I error. Therefore, larger sample size and a more controlled experiment should be conducted to get more insights in the obtained results.

6. Conclusions and implications

This study reports on the effect of reviewers' expertise on design understanding – a constituent part of a design review – using DI and VR. The results show that participants in DI developed higher design understanding than their VR counterparts. However, these findings might also attribute to experience in using a particular technology (VR or CAD). Moreover, groups with higher ME experience and a higher number of ME projects also exhibited higher design understanding, regardless of the environment. Next, although higher contextual experience resulted in increased design understanding for some variables, the overall findings remain inconclusive as other contextual experience variables did not show a similar increasing trend. Finally, intrinsic-dynamic spatial ability did not have a significant effect on design understanding. Similar to the effect of environment, expertise variables might also attribute their effects to other interrelated variables.

These findings have important implications for both research and practice. First, researchers should analyse reviewers' expertise while studying design reviews, even when students perform the review. Only then researchers would be able to identify the subtle differences between VR and DI in design review. Second, we suggest that before a company can assess the usability of VR for their purposes, they should conduct a longer training on the technology and use participants with comparable expertise.

Besides addressing the limitations, future research should examine the effect of expertise levels on design understanding, find a connection between design understanding and design review effectiveness, combine expertise variables and identify which of them have the highest effect.

Acknowledgement

References

Berg, L.P. and Vance, J.M. (2016), "An Industry Case Study: Investigating Early Design Decision Making in Virtual Reality", *Journal of Computing and Information Science in Engineering*, Vol. 17 No. 1, p. 011001. https://doi.org/10.1115/1.4034267

de Casenave, L. and Lugo, J.E. (2017), "Design Review Using Virtual Reality Enabled CAD", *Volume 1: 37th Computers and Information in Engineering Conference*, Cleveland, Ohio, USA, August 6-9, 2017, The American Society of Mechanical Engineers, p. V001T02A067. https://doi.org/10.1115/DETC2017-67878

de Casenave, L. and Lugo, J.E. (2018), "Effects of Immersion on Virtual Reality Prototype Design Reviews of Mechanical Assemblies", *Volume 7: 30th International Conference on Design Theory and Methodology*, Quebec City, Canada, August 26-29, 2018, The American Society of Mechanical Engineers, p. V007T06A044. https://doi.org/10.1115/DETC2018-85542

Coburn, J.Q., Freeman, I. and 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*, Vol. 17 No. 3, p. 031013. https://doi.org/10.1115/1.4036921

Cross, N. (2004), "Expertise in design: an overview", *Design Studies*, Vol. 25 No. 5, pp. 427-441. https://doi.org/10.1016/j.destud.2004.06.002

D'Astous, P. et al. (2004), "Changing our view on design evaluation meetings methodology: a study of software technical review meetings", *Design Studies*, Vol. 25 No. 6, pp. 625-655. https://doi.org/10.1016/j.destud.2003.12.002

Di Marco, M.K., Alin, P. and Taylor, J.E. (2012), "Exploring Negotiation through Boundary Objects in Global Design Project Networks", *Project Management Journal*, Vol. 43 No. 3, pp. 24-39. https://doi.org/10.1002/pmj.21273

Freeman, I.J., Salmon, J.L. and Coburn, J.Q. (2018), "A bi-directional interface for improved interaction with engineering models in virtual reality design reviews", *International Journal on Interactive Design and Manufacturing*, Vol. 12 No. 2, pp. 549-560. https://doi.org/10.1007/s12008-017-0413-0

Gero, J.S. and Kannengiesser, U. (2004), "The situated function-behaviour-structure framework", *Design Studies*, Vol. 25 No. 4, pp. 373-391. https://doi.org/10.1016/j.destud.2003.10.010

Glass, G.V., Peckham, P.D. and Sanders, J.R. (1972), "Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance", *Review of Educational Research*, Vol. 42 No. 3, pp. 237-288. https://doi.org/10.3102/00346543042003237

Goldstone, R.L., Kersten, A. and Carvalho, P.F. (2018), "Categorization and Concepts", In: Thompson-Schill, S.L. and Wixted, J.T. (Eds.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, Volume 3: Language & Thought*, John Wiley & Sons, Hoboken, New Jersey, USA, pp. 275-317. https://doi.org/10.1002/9781119170174.epcn308

Hannah, R., Joshi, S. and Summers, J.D. (2012), "A user study of interpretability of engineering design representations", *Journal of Engineering Design*, Vol. 23 No. 6, pp. 443-468. https://doi.org/10.1080/09544828.2011.615302

Horvat, N. et al. (2019), "Comparing Virtual Reality and Desktop Interface for Reviewing 3D CAD Models", *Proceedings of the Design Society: International Conference on Engineering Design*, Cambridge University Press, Cambridge, Vol. 1, No. 1, pp. 1923-1932. https://doi.org/10.1017/dsi.2019.198

Hu, W.-L. and Reid, T. (2018), "The Effects of Designers' Contextual Experience on the Ideation Process and Design Outcomes", *Journal of Mechanical Design*, Vol. 140 No. 10, p. 101101. https://doi.org/10.1115/1.4040625

Hubka, V. and Eder, W.E. (1988), *Theory of Technical Systems: A Total Concept Theory for Engineering Design*, Springer, Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-642-52121-8

Huet, G. et al. (2007), "Making sense of engineering design review activities", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 21 No. 03, pp. 243-266. https://doi.org/10.1017/S0890060407000261

Lawson, B. and Dorst, K. (2009), *Design Expertise*, Taylor & Francis, London. https://doi.org/10.4324/9781315072043

Liu, Y., Messner, J.I. and Leicht, R.M. (2018), "A process model for usability and maintainability design reviews", *Architectural Engineering and Design Management*, Vol. 14 No. 6, pp. 457-469. https://doi.org/10.1080/17452007.2018.1512042

Morkos, B. and Summers, J.D. (2013), "A study of designer familiarity with product and user during requirement elicitation", *International Journal of Computer Aided Engineering and Technology*, Vol. 5 No. 2/3, p. 139. https://doi.org/10.1504/IJCAET.2013.052934

Mosely, G., Wright, N. and Wrigley, C. (2018), "Facilitating design thinking: A comparison of design expertise", *Thinking Skills and Creativity*, Vol. 27, pp. 177-189. https://doi.org/10.1016/j.tsc.2018.02.004

Newcombe, N.S. and 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*, Springer, Dordrecht, pp. 179-192. https://doi.org/10.1007/978-94-017-9297-4_10

Paes, D., Arantes, E. and 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*, Vol. 84 No. October, pp. 292-303. https://doi.org/10.1016/j.autcon.2017.09.016

Peters, M. et al. (1995), "A Redrawn Vandenberg and Kuse Mental Rotations Test - Different Versions and Factors That Affect Performance", *Brain and Cognition*, Vol. 28 No. 1, pp. 39-58. https://doi.org/10.1006/brcg.1995.1032

Sater-Black, K. and Iversen, N. (1994), "How to conduct a design review", *Mechanical Engineering*, Vol. 116 No. 3, pp. 89-92.

Satter, K. and Butler, A. (2015), "Competitive Usability Analysis of Immersive Virtual Environments in Engineering Design Review", *Journal of Computing and Information Science in Engineering*, Vol. 15 No. 3, p. 031001. https://doi.org/10.1115/1.4029750

Thimmaiah, S., Phelan, K. and Summers, J.D. (2016), "An Experimental Study on the Influence That Failure Number, Specialization, and Controls Have on Confidence in Predicting System Failures", *Journal of Mechanical Design*, Vol. 139 No. 1, p. 011102. https://doi.org/10.1115/1.4034789

Wetmore, W.R., Summers, J.D. and Greenstein, J.S. (2010), "Experimental study of influence of group familiarity and information sharing on design review effectiveness", *Journal of Engineering Design*, Vol. 21 No. 1, pp. 111-126. https://doi.org/10.1080/09544820802238217

Wolfartsberger, J. (2019), "Analyzing the potential of Virtual Reality for engineering design review", *Automation in Construction*, Vol. 104 No. March, pp. 27-37. https://doi.org/10.1016/j.autcon.2019.03.018