VR IN NURSING EDUCATION
- HIGH-FIDELITY VS LOW-FIDELITY IMPLEMENTATIONS

Roger Seiler, Benjamin Lemcke and Andreas Block
Zurich University of Applied Sciences, St. Georgen-Platz 2, P.O. Box, 8401 Winterthur, Switzerland

ABSTRACT
With nursing personnel in short supply, innovative approaches to teaching such as the use of virtual reality (VR) can be one way to attract more students to this field. The beneficial effects of VR in terms of customer motivation, satisfaction, and loyalty suggest that the application of VR software can be affordable in learning settings. However, the implementation of VR software can be expensive. This paper compares a high-interaction fidelity approach (complex and costly implementation) and a low-fidelity approach (a simple and less costly prototype). An experiment with nursing students was conducted to compare the two approaches regarding their effect on self-efficacy, perceived ease of use, perceived usefulness, trust, and intention to use. No relevant or significant differences between high- vs low-interaction fidelity implementation were found. Therefore, we recommend low-fidelity approaches to VR use in delivering basic nursing education as they are easier and less costly to implement without having a negative effect on other relevant constructs.

KEYWORDS
Virtual reality (VR), Nursing Education, High Fidelity, Low Fidelity, Technology Acceptance Model (TAM), Trust

1. INTRODUCTION
In 2020, 153,000 people worked as nursing staff in Switzerland (BFS, 2020). However, there is a nursing staff shortage, and fewer and fewer people are prepared to study, and stay in, nursing. A serious shortage of qualified personnel in Switzerland is forecasted by 2028 (Merçay et al., 2021) while, on a global scale, 5.7 million additional nurses will be needed by 2030 (World Health Organization, 2020). To make matters worse, many graduates, that changed their mind about the profession due to hands on experience gained during their studies, do not transition to the labour market (Merçay et al., 2021), graduation rates are low, and dropout rates high (Dolder and Grünig, 2016).

This paper addresses the issue by offering an approach to enhancing the value of nursing education. Using virtual reality (VR) applications may boost the appeal of nursing programmes and presumably can help to attract and retain more personnel in their profession because VR tends to foster emotions and especially as intrinsic motivation to use VR is high among students (Makransky and Lilleholt, 2018).

Furthermore, VR can also foster intrinsic motivation (Makransky and Lilleholt, 2018) and satisfaction with the education program (Hudson et al., 2019; Makransky and Lilleholt, 2018). The effects are even stronger if a form of delight accompanies satisfaction (Ahrholdt et al., 2016), and VR can be a delightful experience as empirical evidence shows (Potter et al., 2016). In a study on virtual environments (VE), interaction with and immersion in VR also had positive effects on loyalty (Hudson et al., 2019). The element of satisfaction is an important one as 33 percent of nursing staff admitted to being somewhat or completely exhausted in a Swiss study on work-related stress (SECO, 2010). Finally, according to customer management theory, satisfaction is linked to loyalty (Schirmer et al., 2018). It is therefore likely that VR may contribute towards reducing the current nurse shortage.

This paper adds to the existing knowledge of how to implement VR applications in the context of nursing education. In particular, the experiment we conducted sheds light on whether high- or low-interaction fidelity is to be preferred in the context of nursing education. A high interaction fidelity for opening a window is going through each and every step of opening a window (e.g. approaching the window, unlocking it and eventually
opening it and stepping away from the window just opened. A low interaction fidelity example of opening the window could be implemented by just point and clicking on the window and the window is opened with one click. The detailed individual steps needed in between are skipped. In the following, an overview of existing literature on the topic is given.

2. LITERATURE REVIEW

VR has proven to be an effective learning tool. It can provide multisensory stimulation (Chavez and Bayona, 2018), and it can improve social and creative skills (Papanastasiou et al., 2019) as well as memory. The positive effects of VR on learning extend to medical training and education (Cohen et al., 2005). In addition, VR training can be cost effective (Abersold et al., 2012) and improve both procedural knowledge (Dubovi et al., 2017) and performance (Smith and Hamilton, 2015).

Jones et al. (2019) reported positive effects of VR in training midwives. There are prototypes and concepts for visualizing anatomy (Izard and Mendoza, 2016), the emergency training of team leaders of advanced life support (ALS) teams (Moore et al., 2019), and the training of ergonomic transfer of patients (Durr et al., 2021).

Some studies have found that VR contributes to a higher degree to the acquisition of knowledge than traditional methods but not to learning performance (Chen et al., 2020). There are also caveats and pitfalls to using VR in educational settings, for example in tertiary education. According to Evans (2019), these include: 1) the materiality of VR headsets and cables, 2) interfaces and issues with haptics, 3) the ‘language of VR’ to communicate the benefits of VR, 4) cybersickness and gender issues in VR use, and the 5) cost of VR.

Self-efficacy (SE), a term coined by Bandura (1982), refers to trust in one’s own skills – an important concept in the context of human motivation and behaviour (Nissim and Weissbluth, 2017). Self-efficacy has a positive effect on motivation (van Dinther et al., 2011), learning outcomes (Renganayagalu et al., 2019), and, ultimately, students’ performance (Bartimote-Aufflick et al., 2016). This also holds true in the context of nursing education (Zengin et al., 2014).

Research has suggested that VR can have a positive effect on self-efficacy (Nissim and Weissbluth, 2017). There is also a link between display fidelity (Buttussi and Chittaro, 2018; McMahan, 2011), simulation fidelity (Renganayagalu et al., 2019), interaction fidelity (McMahan et al., 2012), and self-efficacy. Interaction fidelity (IF) is essential as well as important concept in the context of learning as immersion and interaction are necessary in order for VR to have a positive effect on learning in terms of learning outcomes, realistic experience, and fostering intrinsic motivation (Chavez and Bayona, 2018).

VR can be seen as technologically sophisticated way of delivering educational content. Whenever technology and users interact, the technology acceptance model (TAM) proposed by (Davis, 1989) can be used, according to which, perceived usefulness (PU), perceived ease of use (PEOU), and intention to use (ITU) play a role in users’ acceptance of technology (Davis, 1985).

When users interact with technology, they become more experienced and develop trust. Research on online shopping has identified a link between trust and the TAM (Gefen et al., 2003) and extended it to VR, which resulted in the VR hardware acceptance model (VR HAM) (Manis and Choi, 2019). However, there is still little research on the TAM and VR in general, let alone on the specific context of nursing staff education. We choose to use this model, however, as it allows us to compare past and more recent research findings and their results. The concept of trust is included because it is relevant to business success (Batsaikhan, 2017) and making purchasing decisions (Hajli et al., 2017).

Although there is research on the effects of VR on self-efficacy, there is a research gap regarding interaction fidelity. Buttussi and Chittaro (2018) focussed entirely on the degrees of freedom as interaction fidelity and in their call for further research suggested taking a closer look at the procedural knowledge gain. Furthermore, Renganayagalu et al. (2019) clearly pointed out that rudimentary VR knowledge of their participants was a limitation and hence complex, high fidelity interaction may be too complicated for novice users.

In our research, IF was modelled as an independent variable and Hypothesis 1 (an overview of the hypotheses and results is given in Table 3) The hypothesis that IF positively influences SE is derived. McMahan et al. (2012) reported a positive effect of interaction fidelity on the usability of the application. Hypothesis 2, therefore, states that IF positively affects PEOU. In line with the TAM, Hypothesis 3 states that IF positively affects PU. No direct effect of IF on trust was found in the existing literature. Nevertheless, simulation fidelity (McMahan et al., 2012) affects PU and PEOU affects trust (Gefen et al., 2003). Hypothesis 4 states that IF
positively affects trust. Hypotheses 5 to 9 are derived from the TAM. As mentioned above, they are listed in Table 3 for the sake of completeness and to test if the TAM holds true in a nursing educational context.

Based on these considerations, the research question is as follows: How should a VR application be implemented in a nursing education context in order to provide value to students and education institutions? There are also four subordinate research questions: What role does interaction fidelity play (SRQ1)? How does interaction fidelity affect technology acceptance (SRQ2)? How is self-efficacy affected (SRQ3)? How is trust affected (SRQ4)?

3. METHOD

Our research using the model of sequential mixed-methods-research (Halcomb and Hickman, 2015) containing a qualitative as well as subsequent quantitative part consists of three phases: 1) expectations and requirements (focus group consisting of seven nursing students), 2) development of the VR prototype, and 3) experiment with the VR prototype (nine students in the treatment group and nine in the control group).

The prototype was developed in Unity (see Figure 1) for the Oculus quest HMD. The treatment group had the high-interaction fidelity version of the VR application, whereas the control group tested the low-interaction fidelity version. The Unity packages XR Plugin Management (V. 4.2.0) and OpenXR Plugin (V. 1.3.1) were used in the prototype. Furthermore, assets from the unity asset store (realistic hands, hospital patient character, hospital room, hospital modular building, props and characters, surgical instruments – hospital props), from cgtrader (IV blood bag dropper stand low-poly 3D model), and from Sketchfab (AZ vaccine vial) were used to implement the scenario. The sound files used (confirm jingle, confirmation downward, wrong, completed, crowd noise light 3_1night) in the prototype are from freesound.org. All sounds were normalised to -1.0 dB in the open-source software Audacity.

Figure 1. Screenshot of the prototype in Unity

The operationalization was implemented as follows: Age, working experience in years, and gender composed the demographic variables (own items). Furthermore, the main constructs of this project were derived from existing literature and validated scales. These were self-efficacy (Riggs et al., 1994), PEOU (Kolitz, 2008), PU (Kolitz, 2008), and trust (Chao, 2019) (see Table 1).
The experiment was conducted in three stages. 1) Participants signed a letter of acceptance confirming that they were taking part voluntarily. 2) They completed a first questionnaire followed by a short instructional video before putting on the Oculus head-mounted display (HMD). A quick check was conducted to ensure that the HMD fitted correctly, and the VR application could be seen clearly. 3) After completing the steps in the VR scenario, participants filled out a second questionnaire. The survey consisted of paper-and-pencil questionnaires, and the software used for analysis was IBM SPSS Statistic 28.0.1.1.

The first phase (focus groups) showed that there is agreement regarding the importance of what activities should be trained using VR. It was generally felt that these should be processes.

In the second phase, a decision was made on what process activity should be implemented by comparing multiple scenarios regarding the value of a VR implementation as well as suitability for VR implementation.

It was found that a suitable scenario would be drawing medicine from a vial and subsequently injecting it into a patient. This scenario consists of the following steps: closing the window, disinfecting the work area, hand disinfection, putting on gloves, removing the cap of the infusion flask, turning the vial upside down 10 times, disinfecting the vial, removing the cap of the syringe, ventilating the vial, drawing the content of the vial into a syringe, injecting the content of the syringe into the infusion flask, disposing of the syringe, applying a sticker to the infusion flask, removing the gloves, docking the infusion flask to the infusion set. The construct “interaction fidelity” was calculated using a framework for interaction fidelity analysis (FIFA) as described in McMahan (2011) for all the activities listed. The calculated FIFA scores were 1.62 (low-interaction fidelity) and 2.79 (high-interaction fidelity).

The third phase involved exposing participants to one of the two (high- vs. low-interaction fidelity) VR scenarios before asking them to fill out the questionnaire.

Of the 18 participants 14 were female and four were male. Compared to the 84.7 percent of women in Switzerland’s nursing workforce (Merçay et al., 2021), the 77 percent in this sample are only seven percent lower. The average age of participants was 21.61 (M = 21.61, SD = 1.539) and their average years of work experience were 1.89 (M = 1.89, SD = 1.875).

### 4. RESULTS, DISCUSSION, FURTHER RESEARCH, AND LIMITATIONS

A first look at the mean values shows that all constructs (SE, PEOU, PU, and trust) have higher mean values in the low-interaction fidelity and high-interaction fidelity versions than in the baseline questionnaire (see Table 2).

Before continuing further, the analysis reliability of the constructs was tested with Cronbach’s Alpha. The reliability is as follows: self-efficacy ($\alpha = .770$), PEOU ($\alpha = .702$), PU ($\alpha = .433$), and trust ($\alpha = .743$) for the baseline questionnaire. The low score of PU may be due to the usefulness was not yet being clear before exposure to the VR application. The reliability of the second questionnaire is as follows: self-efficacy ($\alpha = .606$), PEOU ($\alpha = .924$), PU ($\alpha = .918$), and trust ($\alpha = .728$). If Item 5 (“I have all the necessary skills to do my job.”) in the self-efficacy construct is removed, the reliability falls within the cut-off value proposed by Nunally (1978). The low score of Item 5 can be explained by the fact that participants were students and was, therefore, removed from further analysis.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Variable name (scale)</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Age in years (interval)</td>
<td>own item</td>
</tr>
<tr>
<td></td>
<td>Working experience in years (interval)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender, postcode, hours spent in VR (nominal)</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>7 items (5-point Likert)</td>
<td>(Riggs et al., 1994)</td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>3 items (5-point Likert)</td>
<td>(Kolitz, 2008)</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>4 items (5-point Likert)</td>
<td>(Kolitz, 2008)</td>
</tr>
<tr>
<td>Trust</td>
<td>5 items (5-point Likert)</td>
<td>(Chao, 2019)</td>
</tr>
</tbody>
</table>

Table 1. Overview of operationalisation
A comparison of the high-interaction and low-interaction versions (after exposure to VR application) with the baseline (before exposure to the VR application) shows that, apart from trust, all construct means are higher for the low-interaction fidelity version.

Table 2. Comparison of pre- and post-VR-application exposure questionnaire

<table>
<thead>
<tr>
<th>Construct</th>
<th>Baseline</th>
<th>Low-Interaction Fidelity</th>
<th>High-Interaction Fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>SE</td>
<td>18</td>
<td>2.00</td>
<td>3.89</td>
</tr>
<tr>
<td>PEOU</td>
<td>18</td>
<td>2.67</td>
<td>4.33</td>
</tr>
<tr>
<td>PU</td>
<td>18</td>
<td>3.00</td>
<td>4.67</td>
</tr>
<tr>
<td>Trust</td>
<td>18</td>
<td>3.00</td>
<td>4.50</td>
</tr>
</tbody>
</table>

To test for the effect that interaction fidelity has on the constructs, a Mann-Whitney-U test on high- vs. low-interaction fidelity was conducted as the sample size was rather small and normal distribution could not be assumed. The test had the following values for self-efficacy: U(18) = 20.00, z = -1.82, p < .068, for PEOU U(18) = 40.00, z = -0.05, p < .962, for PU U(18) = 35.00, z = .49, p < .622, and for trust U(18) = 48.00, z = .69, p < .493. None of the constructs show significant differences between high- and low-interaction fidelity.

The conceptual model containing the TAM was verified using correlation analysis (see Figure 2) and, therefore, all the hypotheses are tested (see Table 3). Significant correlations could be reported for H5 and H6 but not for H7. H8 and H9 were not tested as participants had to use the VR application and, hence, intention to use was not tested.

Table 3. Overview of Hypotheses 1 - 9

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Path</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Interaction fidelity → Self-efficacy</td>
<td>Rejected</td>
</tr>
<tr>
<td>H2</td>
<td>Interaction fidelity → Perceived ease of use</td>
<td>Rejected</td>
</tr>
<tr>
<td>H3</td>
<td>Interaction fidelity → Perceived usefulness</td>
<td>Rejected</td>
</tr>
<tr>
<td>H4</td>
<td>Interaction fidelity → Trust</td>
<td>Rejected</td>
</tr>
<tr>
<td>H5</td>
<td>Perceived ease of use → Trust</td>
<td>Accepted</td>
</tr>
<tr>
<td>H6</td>
<td>Perceived ease of use → Perceived usefulness</td>
<td>Accepted</td>
</tr>
<tr>
<td>H7</td>
<td>Trust → Perceived usefulness</td>
<td>Accepted</td>
</tr>
<tr>
<td>H8</td>
<td>Trust → Intention to use</td>
<td>N.a. (compulsory use)</td>
</tr>
<tr>
<td>H9</td>
<td>Perceived usefulness → Intention to use</td>
<td>N.a. (compulsory use)</td>
</tr>
</tbody>
</table>
4.1 Further Research

In the future, researchers might take a closer look at subject loyalty and retention, which are not covered in this preliminary study of high- vs. low-interaction fidelity. Because trust is such a key concept in business and human behaviour and both in the low- and high-interaction fidelity scenarios, and because the mean values (see Table 2) of trust are higher than in the baseline questionnaire, further research may want to focus on the effects that trust has in VR applications in a nursing educational context.

4.2 Limitations

The sample size of our study was rather small, and the average age of our subjects was low because they were students. Gender is biased towards female participants, which is not a real issue given that 84.7 percent of nurses in Switzerland are female. Furthermore, participants asked more questions about the correct use of the VR application after completing the high-interaction fidelity scenario than the low-interaction fidelity scenario. Hence, participants have more trouble interacting with the high-interaction fidelity version and this potentially could have affected the results presented in this paper. The low-reliability score of some constructs must also be mentioned, but as this concerns the baseline questionnaire, it can be explained by participants not being able to answer the questions to PU as they had not yet been exposed to the VR application.

It should also be considered that, coincidentally, participants had another VR experience prior to taking part in this experiment. However, a comparison between high- and low-interaction fidelity is still feasible as the earlier VR experience was not related to this research focus.

Finally, the results presented in this paper hold true for first early-stage projects, early-stage proof-of-concepts and early-stage use cases as the students were not repeatedly exposed to the VR application. Hence, if VR applications are used over a longer period, such as in regular training sessions for a whole semester of study, the results may be different.

5. CONCLUDING REMARKS

The results of our preliminary study apply to activities with a process character (i.e., processes) as the example presented in this paper (injecting the content of a vial into an infusion flask and infusion set). Interaction fidelity does not have a significant direct effect on self-efficacy, perceived ease of use, perceived usefulness, or trust. This result leads to the recommendation that in a nursing education context, low-interaction fidelity implementation is sufficient for activities that have a sequential, process-like character. However, this may not hold true if the interaction with the VR application is prolonged, such as over a whole semester, and if students are repeatedly and regularly exposed to the VR application.

We recommend that educational institutions should take the initial steps towards introducing low-interaction fidelity VR to their nursing programmes, especially for activities with a process character. This is relevant as low-interaction VR applications are less costly to implement, easier to handle, and tend to require fewer explanations than high-interaction ones. In addition, a lower cognitive load imposed on students due to technical aspects leaves more cognitive power for learning and practice.

The (already widely accepted) TAM seems viable in the context of nursing education too. Perceived ease of use affects both perceived usefulness and trust. What is more, both trust and perceived usefulness can lead to word-of-mouth (WOM) dissemination. As a result of relying on online information and referrals (common among the target group, i.e., potential nursing students), WOM can lead to higher student numbers in nursing programmes, which, in turn, can lower acquisition costs.

Usability is a key aspect of VR applications. TAM is, therefore, relevant, in particular with regard to perceived ease of use and trust. This should be kept in mind when implementing low-interaction fidelity VR applications in nursing programmes. With its potential to cut costs and simplify teaching, low-interaction fidelity may lead to the neglect of usability. This is not advised as this may eliminate the potential benefits (for example higher learning performance, higher self-efficacy, higher learning, or higher motivation) VR applications can have.
REFERENCES


Davis, F.D., (1985). A technology acceptance model for empirically testing new end-user information systems.


