Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training

Li, Meng; Ganni, Sandeep; Ponten, Jeroen; Albayrak, Armagan; Rutkowski, Anne-F; Jakimowicz, J.J.

DOI
10.1109/VR46266.2020.1581301697128

Publication date
2020

Document Version
Accepted author manuscript

Published in
Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020

Citation (APA)
Li, M., Ganni, S., Ponten, J., Albayrak, A., Rutkowski, A.-F., & Jakimowicz, J. J. (2020). Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training. In *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020* (pp. 566-572). Article 9089445 (Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020). IEEE. https://doi.org/10.1109/VR46266.2020.1581301697128

Analysing usability and presence of a virtual reality operating room (VOR) simulator during laparoscopic surgery training

Meng Li*
TU Delft IDE
Xi'an Jiaotong University

Sandeep Ganni†
TU Delft IDE
GSL Medical College

Jeroen Ponten‡
Catharina Hospital

Armagan Albayrak §
TU Delft IDE

Anne-F. Rutkowski ¶
Tilburg University

Jack Jakimowicz‖
TU Delft IDE
Catharina Hospital

Figure 1: The concept of a virtual reality operating room

Abstract

Immersive Virtual Reality (VR) laparoscopy simulation is emerging to enhance the attractiveness and realism of surgical procedural training. This study analyses the usability and presence of a Virtual Operating Room (VOR) setup via user evaluation and sets out the key elements for an immersive environment during a laparoscopic procedural training.

In the VOR setup, a VR headset displayed a 360-degree computer-generated Operating Room (OR) around a VR laparoscopic simulator during laparoscopy procedures. Thirty-seven surgeons and surgical trainees performed the complete cholecystectomy task in the VOR. Questionnaires (i.e., Localized Postural Discomfort scale, Questionnaire for Intuitive Use, NASA-Task Load Index, and Presence Questionnaire) followed by a semi-structured interview were used to collect the data.

The participants could intuitively adapt to the VOR and were satisfied when performing their tasks (M=3.90, IQR=0.70). The participants, particularly surgical trainees, were highly engaged to accomplish the task. Despite the higher mental workload on four subscales ($p < 0.05$), the surgical trainees had a lower effort of learning (4 vs 3.33, $p < 0.05$) compared to surgeons. The participants experienced very slight discomfort in seven body segments (0.59-1.16). In addition, they expected improvements for team interaction and personalized experience within the setup.

The VOR showed potential to become a useful tool in providing immersive training during laparoscopy procedure simulation based on the usability and presence noted in the study. Future developments of user interfaces, VOR environment, team interaction and personalization should result in improvements of the system.

Keywords: Laparoscopy simulation, Virtual reality operating room, Surgical training, Presence, Usability, User evaluation

Index Terms: Human-centered computing [Virtual Reality]—Human computer interaction—User evaluation—; Human-centered computing—[Applied Computing]—Life and medical science

1 Introduction

Laparoscopic surgery, also known as minimally invasive surgery (MIS) or keyhole surgery, is a surgical procedure which allows surgeons access to the inside of the body cavity without making a large incision in the skin. This technique has obvious advantages over open surgery, as patients experience less pain and bleeding, a shorter hospital stay and quicker recovery. Laparoscopic surgery is undergoing a fast development and is becoming a standard treatment for many surgical therapies, e.g. cholecystectomy (gallbladder removal surgery) [39]. Robotic surgery is among the latest advances in the laparoscopy field.

Nevertheless, the skills required to perform laparoscopic surgery are largely different from open surgery. During laparoscopic procedures, the surgeons must perform with movements that are more restricted and must work with a narrower field of vision. They must acquire proficiency on non-intuitive motor skills and hand-eye coordination, as well as deal with the ever-changing instruments throughout the procedure [26, 37]. Thanks to the introduction of virtual reality (VR) surgical simulators, the surgeons are able to improve laparoscopic skills without subjecting the patients to unnecessary risk or pain during this learning process [35]. Many reasons along with psychomotor skill and procedural knowledge influence the performance and the mental well-being of surgeons in the operating room (OR) [42]. Research has revealed that distractions are common in the OR and have obvious negative impacts on surgeons'

*e-mail: m.li-4@tudelft.nl
†e-mail: sandeepganni@outlook.com
‡e-mail: jeroenponten@gmail.com
§e-mail: A.Albayrak@tudelft.nl
¶e-mail: a.rutkowski@uvt.nl
‖e-mail: jakimowi@planet.nl

performances and emotions [29, 31]. Hence, training surgeons to handle these challenges requires equally advanced tools that replicate the actual intraoperative distractions.

2 Related works

2.1 VR laparoscopy training

Virtual Reality laparoscopy (VRL) simulation, replicating haptic feedback during procedure-specific tasks, has been proven to accelerate the acquisition of skills of laparoscopic trainees [6]. The main drawback of current VRL simulation is the lack of true representation of the operating theatre experience [17]. Most VRL simulators use a 2D display interface that replicates the tasks but not the environment of busy and often chaotic operating theatres [22, 41]. Numerous distractions occurring in a surgical surrounding, which have been identified and broadly classified into equipment factors, environmental factors, social factors and organizational factors [29]. These distractions increase the task demand and stress level of the surgeons. As Mentis et al stated, residents should be trained both to achieve proficiency and to exercise self-management with distractions in an Operating Room (OR) [23]. Immersive training, representing distraction factors that closely mimic the clinical practice, helps surgical trainees to adapt effectively to their work environments [32].

2.2 VR operating room simulation

To create such a surrounding, the required amount of spatial, financial, personal and technological resources is demanding and can hardly fit into daily clinical routines [1, 17]. Since the upsurge of high-end VR headsets in 2016, it became accessible and affordable to virtually generate an immersive environment of an OR. That environment reproduces distractions as well as generates a good sense of presence, meaning the perception of "being there" in a real OR [15, 24, 34]. Clinical pilot studies have investigated several immersive VR laparoscopic simulators, revealing the face validity and the users' preference of these setups [13, 14, 34]. As no differences in performances appeared between immersive and regular setups, these studies are limited to apparent usefulness or the preferences relating to these immersive environments. However, a key challenge in developing VR-based surgical simulations is to establish usability and a sense of presence from the surgeon's perspective [18]. This topic rarely has been investigated in previous studies.

2.3 User evaluation of VR simulator

It is essential to analyse usability in virtual environments as this analysis demonstrates how intuitively and proficiently users can utilize a product to achieve their objectives [3]. Additionally, mental workload and ergonomic assessments should be incorporated in the evaluation of new laparoscopic training tools, as laparoscopic surgery involves a higher level of mental and physical stress than the open surgery [2, 5, 38]. In medical device development, user evaluation is a common method to identify the usability issues of current setups and indicate potential improvements in future use [43].

A Virtual Operating Room (VOR) setup connecting a VRL simulator and a VR headset was explored in this study. This study analyses the experience of VOR by surgeons and surgical trainees regarding usability and presence in order to identify its potential benefits and improvement opportunities in laparoscopic procedure training.

3 Materials and Methods

3.1 Participants

Thirty-seven Dutch surgeons and surgical trainees were invited to participate in this study between June and August 2018. All participants voluntarily enrolled in the study and signed informed consents. The hospital ethics committee has approved the study. The inclusion criterion for surgical trainees was their prior experience in laparoscopic simulators or box trainers or real operations. The mean age of the participants (male/female = 22:15) was 32.4 years (SD=11.6). The sample was composed of eight experienced surgeons (more than 200 cases) and twenty-nine residents and trainees (two had 101 to 200 cases; three had 51 to 100 cases; twenty-four had 50 or fewer cases). In this article, we refer to the surgeon as "expert" and to the surgical trainee as "novice". Twelve participants had experience on VR or AR technologies (4 for high-end VR, 4 for cardboard VR, 2 for AR Apps, 4 for simulators).

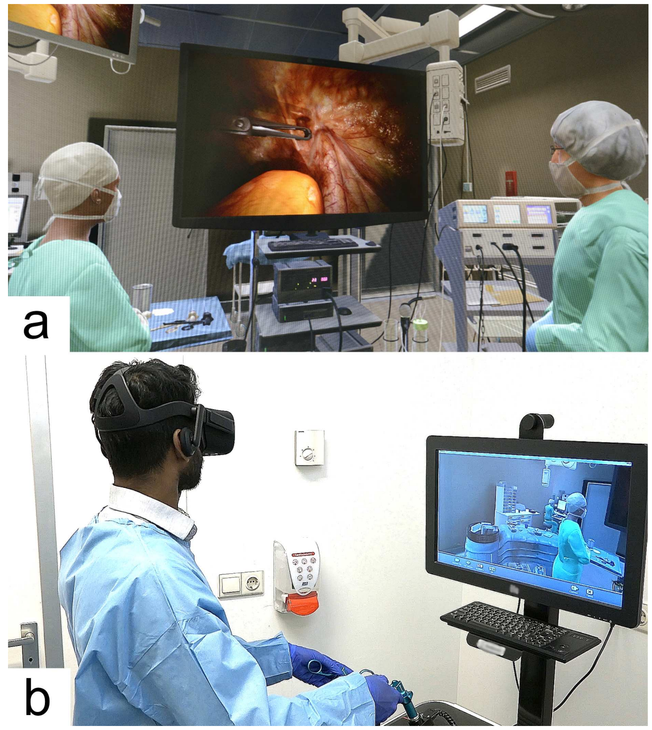

Figure 2: The setup of the VOR simulator. a) The replicated OR surrounding in the VOR and b) An external view of the setup of the VOR simulator.

3.2 Platform

The VOR setup we applied comprised three components: a VR laparoscopic simulator, a VR headset and a virtual OR environment (Figure 1).

The VR laparoscopic simulator was a LapMentor III (Simbionix™, 3D Systems Corporation, USA) with MentorLearn Software. LapMentor III contains two integrated modules: 1) the interface module is an operation table that simulates the patient's abdomen, the trocars, two instruments, a camera, and a double footswitch. The instruments have five DOF and haptic feedback. The footswitch activates electrosurgical coagulation during the training. A freeze mode of the camera allows trainees to navigate it by themselves during operations. The entire module is adjustable in height from 62.99" at the lowest position to 70.86" at the highest. 2) The processing module houses a two-unit industrial PC with a 24" touch-screen monitor (1920*1080 dpi): (a) the simulation unit is a 3.1-GHz Intel Core i7-4770S and an Intel™ Motherboard; (b) the VOR unit is an NVIDIA GeForce GTX 1060 graphic card and an Intel™ SHARKBAY Motherboard. Both units run on Windows 7 Professional (x64) operating system.

The software includes a basic skills trainer and a procedural skills trainer. The basic skills trainer allows trainees to practice tasks that are abstractions of those performed during surgery. The procedural skills trainer is a simulation that allows trainees to perform an entire laparoscopic cholecystectomy with virtual patients. The trainee could see a computer-generated body cavity during operations through the monitor. If trainees want to change tools in LapMentor, they need to: 1) pull out an instrument to see a pop-up menu on the screen, 2) hold and pull the instrument left or right to choose one, and then 3) clip the instrument to select and insert it again.

The VR headset was a 2016 Oculus Rift model, providing stereoscopic images (1080 * 1200 per eye, 110° field of view), integrated 3D audio and 6 DOF head-tracking. The virtual OR was a 360-degree computer-generated environment that replicates a real OR, including a full setup of instruments and equipment and as a new feature, a surgical team and various distractions. The distractions covered three most frequently occurring types: door movements, phones/pagers/bleepers, and radio, as well as one most distracting type: case-related communication (Figure 2a) [23].

The VR headset displays the virtual OR around the simulator while a trainee is practising the cholecystectomy, and a virtual instructor talks to the trainee throughout the procedure (Figure. 2a right-hand side). If the trainee changes a tool in VOR, there are several differences from the LapMentor: 1) the tool menu is floating at eye level; 2) turn a knot at the front of the handle to choose tools instead of pulling the instrument. To simulate the electrosurgical coagulation, a footswitch is displayed underneath the simulated monitor.

3.3 Procedure

Participants performed a task (LapMentor III: complete cholecystectomy) after a standardized introduction from the researchers [4]. Researchers informed participants that the purpose of the study is to investigate the use of VOR in surgical procedural tasks for immersive training. A pre-test protocol limited the time of the task to 15 minutes according to the empirical duration to complete it. After completing the task, participants answered four questionnaires regarding the usability and presence. A semi-structured interview allowed collecting the surgeons' narratives.

In this study, the usability of the VOR was evaluated with a combination of three questionnaires. First, intuitiveness, in other words subconsciously applying prior knowledge, was evaluated via the Questionnaire for Intuitive Use (QUESI) [25]. The QUESI was applied across multiple professions, including healthcare, to quantify intuitiveness of virtual environments [21, 33]. The validated assessment asked if the VOR appears intuitive and satisfying using a 5-point Likert scale (1= fully disagree, 5=fully agree). Second, the mental workload of performing the task in the VOR was measured using the NASA-TLX [11]. This validated tool has already extensively been used for assessing the task demand of surgeons when performing laparoscopic surgeries or training [20, 46]. The participants gave a score to the levels of mental, physical and temporal demands they perceived, as well as their effort, performance and frustration during the task. The Raw Task Load Index (RTLX) and subscales were calculated into a score between 0 and 100 (0=low, 100=high) [11]. Third, to assess the physical stress, perceived as discomfort, we used a validated assessment - Localized Postural Discomfort (LPD) [10, 19]. The participants rated the symptoms of discomfort in every segment of their body via a 10-point scale (0=no discomfort, 10=extreme discomfort). The answers were categorized as insupportable discomfort when participants marked the value as more than 2 according to the ISO/FDIS 11226 [16].

The factors influencing the perception of presence were investigated via a questionnaire followed by an interview. The Presence Questionnaire, a known assessment instrument, was modified and previously validated [44, 45]. In this study, we added two items (i.e. accuracy of gestures, realistic resistance of tissue) on "haptic" and one item on "sound" (realistic sound effect) according to the features of the VOR, and applied a 21-point scale (1= not at all, 21=completely) to survey the presence in fine gradients [7].

Table 1: The level of intuitive use of the VOR (1= "Fully disagree", 5= "Fully agree").

QUESI	Total Mean(IQR)	Novice Mean(IQR)	Expert Mean(IQR)
Low subjective mental workload	3.67 (1.33)	4.00 (1.17)	3.67 (1.17)
High perceived achievement of goals	4.00 (1.33)	4.00 (0.50)	3.67 (1.75)
Low perceived effort of learning	4.00 (1.00)	4.00(0.83)*	3.33 (1.25)
High familiarity	3.67 (0.50)	3.67 (0.50)	3.67 (0.83)
Low perceived error rate	4.00 (0.50)	4.00 (1.00)	3.25 (1.75)
Total	3.90 (0.70)	3.90 (0.53)	3.38 (1.33)

Note:*Statistically significant results with $p<0.05$.

The semi-structured interview consisted of two questions: (1) How satisfied are you with the Virtual OR experience? (2) Which factors were not compelling or not realistic in the Virtual OR experience?

3.4 Statistical Analysis

The data were analysed using SPSS v.25. Descriptive statistics of each questionnaire were calculated, including mean and standard deviation (SD), or median and interquartile range (IQR). The comparison of means used one-sample t-test (normally distributed) or Wilcoxon signed-rank test (non-normally distributed). The differences between novices and experts were tested using a classical independent-sample t-test; otherwise, non-parametric tests such as the Kruskal-Wallis test and the Mann-Whitney U test were utilized where appropriate. A p-value of <0.05 was considered as statistically significant.

4 Results

4.1 Intuitive Use

The participants, at a minimum, agreed (>score 3 "neutral") that the VOR appeared intuitive and satisfying to perform laparoscopic procedural training (M=3.90, IQR=0.70). The perceived achievement of goals (M=4.00, IQR=1.33) and error rate (M=4.00, IQR=0.50) seemed to be the most intuitive factors, related to a highly effective interaction;the perceived effort of learning is also intuitive (M=4.00, IQR=1.00), related to applying prior knowledge for the first-time use. The novices rated four subscales more intuitive than the experts, while the perceived effort of learning was significantly different (4.00 vs 3.33, p <0.05, Mann-Whitney U test) (Table 1).

4.2 Mental Workload

Thirty-seven participants rated the overall mental workload (RTLX Mean=39.96, SD=14.53) lower than the midpoint of the full range (0-100), indicating the VOR imposed a moderate demand on the users. The subscales varied from 51.49 on effort to 27.30 on frustration (Table 2). It seemed that the mental demand (M=52.16, SD=22.66) and effort (M=51.49, SD=19.43), i.e. intellectual work and required proficiency, were the key components of the mental workload in the VOR. The novices had a significant higher workload on mental demand (56.72 vs 35.63, p= .019), physical demand (40.17 vs 20.63, p= .011), temporal demand (37.93 vs 18.13, p= .006), effort (55.34 vs 37.50, p= .019), and overall workload (43.16 vs 28.23, p=.008) than the experts (Mann-Whitney U Test).

Table 2: Self-reported mental workload after training in the VOR (0-100, the higher score means higher mental workload)

NASA-TLX	Total Mean(SD)	Novice Mean(SD)	Expert Mean (SD)
Mental Demand	52.16(22.66)	56.72(20.76)*	35.63(22.75)
Physical Demand	35.95(21.40)	40.17(20.68)*	20.63(17.41)
Temporal Demand	33.65(21.62)	37.93(20.24)*	18.13(20.34)
Performance	39.05(19.03)	40.34(19.36)	34.38(18.21)
Effort	51.49(19.43)	55.34(18.51)*	37.50(16.90)
Frustration	27.30(20.97)	28.45(18.28)	23.13(29.99)
RTLX	39.93(14.53)	43.16(13.10)*	28.23(14.11)

Note:Statistically significant results with $p<0.05$.

Table 3: Localised Postural Discomfort (LPD) of body segments (0="No Discomfort", 10="Extreme Discomfort")

Body segments	Total Mean(SD)	Novice Mean(SD)	Expert Mean (SD)
Neck	0.78 (1.29)	0.79 (1.35)	0.75 (1.16)
Lower neck (L/R)	0.59 (0.93)	0.62 (0.94)	0.50 (0.93)
Hand(L)	1.16 (1.77)	1.34 (1.93)	0.50 (0.76)
Hand(R)	0.70 (1.22)	0.86 (1.33)	0.13 (0.35)
Eye(L/R)	0.97 (1.57)	1.00 (1.56)	0.88 (1.73)

4.3 Comfort

The average discomfort in each body segment ranged from 0.05 to 1.16, corresponding to almost no discomfort to very low discomfort. The scores of seven body segments out of all twenty-three parts (30.4%) were above the slightest discomfort level (score 0.5) (Table 3), while only the left hand had a significantly higher discomfort (1.16 vs 0.5, $p<0.05$, one-sample t-test). In the left hand and both eyes (n=6, 16.2%), as well as the neck (n=7, 18.9%), some participants experienced insupportable discomfort. No significant difference was found between novices and experts regarding the physical comfort ($p > 0.1$, Mann-Whitney U Test).

4.4 Presence

In the VOR, Self-evaluated performance seemed most important to presence, as participants adjusted to the environment very quickly (M=16.39, SD=1.90), and could move and interact proficiently at the end of the task (M=16.17, SD=2.10) (see Table 4 and Supplementary Table). The Sound (M=14.79, SD=2.69) appeared mainly to contribute to presence as well, in that participants could easily recognize and localize sounds and viewed the sounds as realistic. The Quality of interface seemed to facilitate the presence perception the least, and the instrument interface had the lowest rank (M=9.97, SD=3.81, see Supplementary Table). Both novices and experts had similar presence level across the subscales ($p > 0.2$, independent-sample t-test).

Table 4: Average rate on subscales of Presence Questionnaire (1="Not at all", 11="Somewhat", 21="Completely")

Presence	Total Mean(SD)	Novice Mean(SD)	Expert Mean (SD)
Realism	14.02 (2.75)	14.15 (2.95)	13.55 (1.97)
Possibility to act	14.24 (2.42)	14.08 (2.55)	14.84 (1.92)
Quality of interface	11.70 (3.38)	11.64 (3.55)	11.92 (2.95)
Possibility to examine	14.24 (2.70)	14.67 (2.77)	12.75 (1.88)
Self-evaluation of performance	16.28 (2.10)	16.55 (2.13)	15.31 (1.79)
Haptic	13.33 (2.78)	14.82 (2.90)	14.66 (1.88)
Sound	14.79 (2.69)	13.49 (2.82)	12.75 (2.72)

4.5 Interview

Thirty-five participants reported that they felt actually had been present in an OR and were engaged by the scenario. The majority (25/37) of the participants mentioned the talk and the sounds enhanced their presence. The participants, particularly the surgical trainees, were highly engaged and excited to complete the procedure. We broadly categorized participant's narratives on the presence of VOR into user interfaces, VOR environment, team interaction and personalization considering the factors of distractions [29].

4.5.1 User interfaces

Trocar: eight participants, especially surgeons, struggled with many slips and were annoyed by the way of switching instruments. The surgeons and experienced trainees(>100 cases) reported the haptic resistance as too low. A delay in changing instruments was found. *Headset*: especially for people with corrected vision, participants often encountered a problem to see a clear image from one or both eyes. The low graphic resolution was also reported. The participants with eyewear (4 in total) had difficulty to put on the VR headset correctly on top of their glasses. The VR headset could press on the glasses and caused a high level of discomfort or even pain in the face.

4.5.2 VOR environment

OR setup: two participants noted that they could not find the footswitch in the VOR because the feet were missing. The additional factors included the incorrect OR layout, disproportionate elements and unrealistic rendering, e.g. the wrong direction of monitor towards the patient's bed, or the size of the monitor. *Surgery steps*: two participants commented that the procedures of the laparoscopy would vary slightly from case to case, while the steps in the VOR seemed to be more rigid. *Sounds*: three participants stated that the sound seemed too loud considering the space of the VOR.

4.5.3 Team interaction

Instructions: four participants were confused by the repetitive instruction from the avatar when the action had already been performed. *Camera assist*: most participants noticed that teamwork was missing, so they had to lay down the instruments carefully and navigated the camera by themselves. The participants who had real OR experience suggested that an assistant should hold the camera and follow the surgeon's manoeuvre throughout operations. *Mood*: two participants remarked that the communication was impersonal and needed some added emotion. An additional comment was that the team was mainly motionless; in reality, the team would move around, if only slightly.

4.5.4 Personalization

Nine participants said they ignored the instructions as background noise because the other surgeon's name was called. Two surgeons asked for background music that they could switch it on or off. Four surgeons expected communication in their native language.

5 Discussion

Training procedural tasks under immersive virtual contexts are already in widespread use in military and aviation industry [27, 28]. Immersive training simultaneously facilitates the acquisition of technical and non-technical skills (e.g. communication and teamwork) owing to distraction simulation [8]. Creating immersive training in skills labs is crucial in acquiring skills and intellectual abilities to optimize patient safety and preserve surgeons' resources essential to the laparoscopy process [31, 36]. The VOR outlined and evaluated in this study built on the advantages of VR laparoscopy simulation, and integrated the immersive experience of an OR. The results demonstrated clearly that immersive training via a VR headset heightens the motivation of trainees and demonstrated a new dimension to integrate immersive OR context in surgical procedural training. The surgical trainees in most European countries were kept from simulation-based training by various external demotivating factors, such as long working hours, limited free time, the overload of clinical work [17]. It is therefore relevant to develop a training setup to boost and sustain trainee's motivation, which is a key element of a successful delivery of laparoscopy training curricula [17].

5.1 Usability and Presence

The results of the usability questionnaires indicated a good sense of intuitiveness, little physical stress and moderate mental workload when performing tasks in the VOR. The simulation of auditory distractions, such as radio, phone-call, pagers and bleeps, most frequently occurring in the OR, enhanced participant's sense of familiarity as they commented. Auditory distractions might result in increased mental effort on inexperienced trainees, which has been suggested by several experimental studies [23]. In this study, we also found that the mental workload was perceived significantly higher by the novices than the experts. The visual stimuli, such as door movement, either in real OR or in simulated conditions, did not obviously affect the flow of procedures [23, 29]. The mental distraction, i.e. case-related communication, was perceived as less annoying when the participants were highly absorbed [31]. The participants rated mental demand as high and frustration as low, indicating that they tended to enjoy intellectual challenges created by the VOR, which also confirmed by their narratives [32].

As we expected, the novices recognized effort as the main source of mental workload, indicating that the distractions in the VOR influenced their flow of performing. The increased mental workload, triggered by the integrated tasks and distractions, created the condition for novices to perform better in a real work environment. This also has been suggested by a previous study, which investigated the role of distraction and mental load during a VRL simulation [32].

Most factors of presence questionnaire revealed that the VOR was perceived as adequately immersive. The participants were satisfied with their quick adaption and proficiency of interaction in the VOR as shown by the QUESI as well. The sound aspect was compelling to the participants in that they could easily recognize and localize different sounds. In addition, the sound effect was perceived as realistic like in a real OR.

5.2 The improvements of the VOR

The presence questionnaire and the participant's comments pinpoint the user interfaces as the most salient limitation of current VOR. The haptic interface provided accurate feedback on gestures, but less realistic experience on interaction and the resistance of tissues. This appeared to relate to the intuitiveness of the experts and the discomfort in hands. As we observed, the surgeons slowed down and made most of their errors during instrument switching. The surgeons also reported their struggling to adapt to the unnatural way of switching tools. The fatigue on the left hand might attribute to a tight hold of the instrument caused by low fidelity of the haptic feedback, which is well recognized in VR surgical simulators [18]. Considering the visual interface, the participants with corrected vision often experienced insupportable discomfort in the eyes due to the incompatibility of the headset for glasses or contact lens. The discomfort in the neck may be associated with the weight of the headset.

The environmental setup such as OR layout and team placement was viewed to be fundamental for a realistic OR experience. To match the VOR to a real OR, we assume that panoramic video or volumetric video is a promising alternative or complement as it regenerates OR scenarios by filming them in a real OR [12, 15]. These technologies could therefore accurately replicate what happens within a real OR including distraction factors.

The creation of an immersive team interaction largely attributes to mimicking mental distractions happening throughout the surgical procedure. These distractions range from procedural distractions, such as camera manipulation, procedure-related conversations, to social distractions, like case-irrelevant or medical-irrelevant communication. Novices needed a considerable amount of mental resources to construct cognitive schemata of the surgical procedure; and accomplishing tasks with additional distractions required extra mental resources, which is even more demanding [32]. Social distractions, like patient-irrelevant and case-irrelevant conversation, play a role to reduce stress, particularly when the task engagement is high. We may thus infer that introducing a virtual team with better-designed distractions reduces required mental resources and helps novices to concentrate on their flow. In this way, the trainees would accelerate the construction of these schemata [30]. This approach might contribute to the transfer from conscious competences to unconscious competences. As the Crew Resource Management (CRM) strategy is missing in current laparoscopy curricula, the virtual team might offer a potential to integrate CRM into procedural laparoscopy training curricula in the near future [40].

Additionally, the semi-structured interview showed a strong emphasis on user's (surgical trainees and surgeons) needs for personalization. It was viewed as as a main factor to enrich a realistic and immersive experience. Personalization pertaining to instruction and language, instrumentation, and background music is expected to match user's needs, wishes and expectations in a real OR. The potential of customizing the environment should be given some serious thought, taking into account specific demands, related to the region, the country or even the institution where the training takes place.

5.3 Limitations

The outcome of this study demonstrates the effectiveness of a VR-based distractive environment as a whole for laparoscopic procedural training. This explorative analysis has the following limitations that point out chances for future studies. (1) We deliberately avoided comparing the VOR with either regular VR laparoscopic simulators or real cholecystectomies. The next step will involve analysing and comparing experiences in both settings. (2) The current study mainly included self-assessment, while participants possibly over-assess their performance in a new immersive training [9]. Hence, we suggest that future studies may include self-assessment, objective measurements and expert assessment to triangulate the evaluation on the performance. (3) Future studies should also investigate and compare how the different types of distractions would influence usability, presence and performance.

6 Conclusion

The VOR showed potential to become a useful tool in providing immersive training during laparoscopy simulation based on the usability and presence analysed in this study. We suggest four improvements for a higher level of presence: 1) optimize haptic and visual interfaces; 2) create a virtual OR environment applying alternative solutions, such as cinematic technologies; 3) include a virtual team facilitating non-technical skills training and stress-reducing; 4) investigate the needs of the surgeons for personalized training. We believe that these improvements will increase the effectiveness of the VOR for laparoscopy training, increase the motivation and speeding up the process of adaption of the trainees to the real OR setting.

Acknowledgments

The authors thank the staff of the Skills Lab in Catharina Hospital Eindhoven and Doris Aschenbrenner from Delft University of Technology for their support in facilitating this study. Meng Li is funded by a PhD scholarship (201706280020) from the Chinese Scholarship Council to conduct this study.

References

[1] I. Badash, K. Burtt, C. A. Solorzano, and J. N. Carey. Innovations in surgery simulation: a review of past, current and future techniques. *Annals of translational medicine*, 4(23):453–453, 2016. doi: 10.21037/atm.2016.12.24

[2] R. Berguer, W. Smith, and Y. Chung. Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. *Surgical Endoscopy*, 15(10):1204–1207, 2001. doi: 10.1007/s004640080030

[3] D. A. Bowman, J. L. Gabbard, and D. Hix. A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence: Teleoperators and Virtual Environments*, 11(4):404–424, 2002. doi: 10.1162/105474602760204309

[4] S. Buzink, M. Soltes, J. Radonak, A. Fingerhut, G. Hanna, and J. Jakimowicz. Laparoscopic Surgical Skills programme: preliminary evaluation of Grade i Level 1 courses by trainees. *Wideochirurgia i inne techniki maloinwazyjne = Videosurgery and other miniinvasive techniques*, 7(3):188–192, 2012. doi: 10.5114/wiitm.2011.28895

[5] C. M. Carswell, D. Clarke, and W. B. Seales. Assessing mental workload during laparoscopic surgery. *Surgical Innovation*, 12(1):80–90, 2005. doi: 10.1177/155335060501200112

[6] S. R. Dawe, G. N. Pena, J. A. Windsor, J. A. Broeders, P. C. Cregan, P. J. Hewett, and G. J. Maddern. Systematic review of skills transfer after surgical simulation-based training. *Br J Surg*, 101(9):1063–76, 2014. doi: 10.1002/bjs.9482

[7] R. F. Dyer, J. J. Matthews, C. E. Wright, and K. L. Yudowitch. *Questionnaire construction manual*. OPERATIONS RESEARCH ASSOCIATES PALO ALTO CA, July 1976.

[8] R. Flin, P. O'Connor, and K. Mearns. Crew resource management: improving team work in high reliability industries. *Team Performance Management: An International Journal*, 8(3/4):68–78, 2002. doi: 10.1108/13527590210433366

[9] S. Ganni, M. K. Chmarra, R. H. M. Goossens, and J. J. Jakimowicz. Self-assessment in laparoscopic surgical skills training: Is it reliable? *Surgical Endoscopy*, 31(6):2451–2456, 2017. doi: 10.1007/s00464-016-5246-6

[10] H. H. Hamberg-van Reenen, A. J. van der Beek, B. M. Blatter, M. P. van der Grinten, W. van Mechelen, and P. M. Bongers. Does musculoskeletal discomfort at work predict future musculoskeletal pain? *Ergonomics*, 51(5):637–648, 2008. doi: 10.1080/00140130701743433

[11] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909

[12] Z. Huang, T. Li, W. Chen, Y. Zhao, J. Xing, C. LeGendre, L. Luo, C. Ma, and H. Li. Deep volumetric video from very sparse multi-view performance capture. In *Computer Vision – ECCV 2018*, pp. 351–369. Springer, Munich, 2018. doi: 10.1007/978-3-030-01270-0_21

[13] T. Huber, M. Paschold, C. Hansen, H. Lang, and W. Kneist. Artificial versus video-based immersive virtual surroundings: Analysis of performance and user's preference. *Surgical Innovation*, 25(3):280–285, 2018. doi: 10.1177/1553350618761756

[14] T. Huber, M. Paschold, C. Hansen, T. Wunderling, H. Lang, and W. Kneist. New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff. *Surg Endosc*, 31(11):4472–4477, 2017. doi: 10.1007/s00464-017-5500-6

[15] T. Huber, T. Wunderling, M. Paschold, H. Lang, W. Kneist, and C. Hansen. Highly immersive virtual reality laparoscopy simulation: development and future aspects. *International Journal of Computer Assisted Radiology and Surgery*, 13(2):281–290, 2018. doi: 10.1007/s11548-017-1686-2

[16] International Organization for Standardization, Geneva. *ISO/FDIS 11226: Ergonomics–Evaluation of static working postures*, July 2000.

[17] J. J. Jakimowicz and S. Buzink. Training curriculum in minimal access surgery. In *Training in Minimal Access Surgery*, pp. 15–34. Springer, London, 2015. doi: 10.1007/978-1-4471-6494-4_2

[18] A. Koch, M. Pfandler, P. Stefan, P. Wucherer, M. Lazarovici, N. Navab, U. Stumpf, R. Schmidmaier, J. Glaser, and M. Weigl. Say, what is on your mind? surgeons' evaluations of realism and usability of a virtual reality vertebroplasty simulator. *Surgical Innovation*, 26(2):234–243, 2019. doi: 10.1177/1553350618822869

[19] C. P. Kruizinga, N. J. Delleman, and J. M. Schellekens. Prediction of musculoskeletal discomfort in a pick and place task (a pilot study). *Int J Occup Saf Ergon*, 4(3):271–286, 1998. doi: 10.1080/10803548.1998.11076394

[20] G. I. Lee, M. R. Lee, T. Clanton, E. Sutton, A. E. Park, and M. R. Marohn. Comparative assessment of physical and cognitive ergonomics associated with robotic and traditional laparoscopic surgeries. *Surgical Endoscopy*, 28(2):456–465, 2014. doi: 10.1007/s00464-013-3213-z

[21] M. Li, A. Albayrak, Y. Zhang, and D. van Eijk. Multiple factors mental load evaluation on smartphone user interface. In *Proc.IEA 2018*, pp. 302–315. Springer, Florence, 2019. doi: 10.1007/978-3-319-96059-3_33

[22] G. L. McCreery, M. El-Beheiry, and C. M. Schlachta. Local and national laparoscopic skill competitions: residents' opinions and impact on adoption of simulation-based training. *Surgical Endoscopy*, 31(11):4711–4716, 2017. doi: 10.1007/s00464-017-5546-5

[23] H. M. Mentis, A. Chellali, K. Manser, C. G. L. Cao, and S. D. Schwaitzberg. A systematic review of the effect of distraction on surgeon performance: directions for operating room policy and surgical training. *Surgical Endoscopy*, 30(5):1713–1724, 2016. doi: 10.1007/s00464-015-4443-z

[24] D. Mestre, P. Fuchs, A. Berthoz, and J. Vercher. Immersion et présence. *Le traité de la réalité virtuelle. Paris: Ecole des Mines de Paris*, pp. 309–38, 2006.

[25] A. Naumann and J. Hurtienne. Benchmarks for intuitive interaction with mobile devices. In *Proc. MobileHCI'10*, pp. 401–402. ACM, Lisbon, Portugal, 2010. doi: 10.1145/1851600.1851685

[26] I. Nisky, F. Huang, A. Milstein, C. M. Pugh, F. A. Mussa-Ivaldi, and A. Karniel. Perception of stiffness in laparoscopy–the fulcrum effect. *Studies in health technology and informatics*, 173:313, 2012.

[27] J. Osterlund and B. Lawrence. Virtual reality: Avatars in human spaceflight training. *Acta Astronautica*, 71:139–150, 2012. doi: 10.1016/j.actaastro.2011.08.011

[28] N. Pallavicini, F.and Toniazzi, L. Argenton, L. Aceti, and F. Mantovani. Developing effective virtual reality training for military forces and emergency operators: From technology to human factors. In *Proc. MAS 2015*, pp. 206–210. Dime University of Genoa, Bergeggi, Italy, 2015. doi: 10.13140/RG.2.1.3590.4484

[29] M. C. Persoon, H. J. H. P. Broos, J. A. Witjes, A. J. M. Hendrikx, and A. J. J. M. Scherpbier. The effect of distractions in the operating room during endourological procedures. *Surgical endoscopy*, 25(2):437–443, 2011. doi: 10.1007/s00464-010-1186-8

[30] J. Pluyter. *Designing immersive surgical training against information technology-related overload in the operating room*. Thesis, Tilburg University, 2012. doi: urn:nbn:nl:ui:12-5660756

[31] J. R. Pluyter, S. N. Buzink, A.-F. Rutkowski, and J. J. Jakimowicz. Do absorption and realistic distraction influence performance of component task surgical procedure? *Surgical Endoscopy*, 24(4):902–907, 2010. doi: 10.1007/s00464-009-0689-7

[32] J. R. Pluyter, A.-F. Rutkowski, and J. J. Jakimowicz. Immersive training: breaking the bubble and measuring the heat. *Surgical Endoscopy*, 28(5):1545–1554, 2014. doi: 10.1007/s00464-013-3350-4

[33] P. Saalfeld, A. Mewes, M. Luz, B. Preim, and C. Hansen. Comparative evaluation of gesture and touch input for medical software. In *Proc. Mensch und Computer 2015*, pp. 143–152. Oldenbourg Wissenschaftsverlag, Stuttgart, 2015.

[34] G. Sankaranarayanan, B. Li, K. Manser, S. B. Jones, D. B. Jones, S. Schwaitzberg, C. G. L. Cao, and S. De. Face and construct validation of a next generation virtual reality (gen2-vr©) surgical simulator. *Surgical Endoscopy*, 30(3):979–985, 2016. doi: 10.1007/s00464-015-4278-7

[35] M. Schijven, J. Jakimowicz, I. Broeders, and L. Tseng. The Eindhoven laparoscopic cholecystectomy training course—improving operating room performance using virtual reality training: Results from the first EAES accredited virtual reality training curriculum. *Surgical Endoscopy*, 19(9):1220–1226, 2005.

[36] M. P. Schijven and J. J. Jakimowicz. Validation of virtual reality simulators: Key to the successful integration of a novel teaching technology into minimal access surgery. *Minimally Invasive Therapy and Allied Technologies*, 14(4-5):244–246, 2005. doi: 10.1080/13645700500221881

[37] D. J. Scott, P. C. Bergen, R. V. Rege, R. Laycock, S. T. Tesfay, R. J. Valentine, D. M. Euhus, D. R. Jeyarajah, W. M. Thompson, and D. B. Jones. Laparoscopic training on bench models: better and more cost effective than operating room experience? *Journal of the American College of Surgeons*, 191(3):272–283, 2000.

[38] M. Silvennoinen and L. Kuparinen. Usability challenges in surgical simulator training. In *Proc. ITI 2009*, pp. 455–460. IEEE, Dubrovnik, Croatia, 2009. doi: 10.1109/ITI.2009.5196126

[39] S. J. Spaner and G. L. Warnock. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *Journal of Laparoendoscopic and Advanced Surgical Techniques*, 7(6):369–373, 1997.

[40] S. Undre, M. Koutantji, N. Sevdalis, S. Gautama, N. Selvapatt, S. Williams, P. Sains, P. McCulloch, A. Darzi, and C. Vincent. Multidisciplinary crisis simulations: The way forward for training surgical teams. *World Journal of Surgery*, 31(9):1843–1853, 2007. doi: 10.1007/s00268-007-9128-x

[41] K. W. van Dongen, W. A. van der Wal, I. H. M. B. Rinkes, M. P. Schijven, and I. A. M. J. Broeders. Virtual reality training for endoscopic surgery: voluntary or obligatory? *Surgical endoscopy*, 22(3):664–667, 2008. doi: 10.1007/s00464-007-9456-9

[42] A. Vereczkel, H. Bubb, and H. Feussner. Laparoscopic surgery and ergonomics: It's time to think of ourselves as well. *Surgical Endoscopy*, 17(10):1680–1682, 2003.

[43] M. E. Wiklund P.E., J. Kendler, and A. Y. Strochlic. *Usability testing of medical devices*. CRC press, Boca Raton, 2*nd* ed., 2015. doi: 10.1201/b19082

[44] B. G. Witmer, C. J. Jerome, and M. J. Singer. The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3):298–312, 2005. doi: 10.1162/105474605323384654

[45] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998. doi: 10.1162/105474698565686

[46] B. Zheng, X. Jiang, G. Tien, A. Meneghetti, O. N. M. Panton, and M. S. Atkins. Workload assessment of surgeons: correlation between nasa tlx and blinks. *Surgical endoscopy*, 26(10):2746–2750, 2012. doi: DOI 10.1007/s00464-012-2268-6