



Learning procedural skills with a virtual reality simulator: An acceptability study[☆]



Marie-Stéphanie Bracq^{a,d}, Estelle Michinov^{a,*}, Bruno Arnaldi^b, Benoît Caillaud^c, Bernard Gibaud^d, Valérie Gouranton^b, Pierre Jannin^d

^a Univ Rennes, LP3C (EA 1285), F-35000 Rennes, France

^b INSA Rennes, IRISA/Inria, F-35000 Rennes, France

^c Hycomes, INRIA, F-35000 Rennes, France

^d Univ Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

ARTICLE INFO

Keywords:

VR simulator
Procedural skills
Acceptability
Nursing education

ABSTRACT

Background: Virtual Reality (VR) simulation has recently been developed and has improved surgical training. Most VR simulators focus on learning technical skills and few on procedural skills. Studies that evaluated VR simulators focused on feasibility, reliability or easiness of use, but few of them used a specific acceptability measurement tool.

Objectives: The aim of the study was to assess acceptability and usability of a new VR simulator for procedural skill training among scrub nurses, based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model.

Participants: The simulator training system was tested with a convenience sample of 16 non-expert users and 13 expert scrub nurses from the neurosurgery department of a French University Hospital.

Methods: The scenario was designed to train scrub nurses in the preparation of the instrumentation table for a craniotomy in the operating room (OR).

Results: Acceptability of the VR simulator was demonstrated with no significant difference between expert scrub nurses and non-experts. There was no effect of age, gender or expertise. Workload, immersion and simulator sickness were also rated equally by all participants. Most participants stressed its pedagogical interest, fun and realism, but some of them also regretted its lack of visual comfort.

Conclusion: This VR simulator designed to teach surgical procedures can be widely used as a tool in initial or vocational training.

1. Introduction

Virtual reality (VR) technology has rapidly evolved in recent years, with applications in different fields, such as gaming, aviation and military training, education, learning and social skills training, simulations of surgical procedures, and psychological treatments (Cipresso et al., 2018). VR systems enable the development of low-cost, realistic, easy-to-use, easily configurable simulators that reduce safety, ethical and health problems. VR simulation can thus be used in high-risk settings for medical education and surgical training for example, to assist trainees acquire skills in a safe environment (Agha and Fowler, 2015).

Additionally, some VR systems allow automatic recording of actions for quantitative and objective performance evaluation. Studies assessing VR simulators have generally focused on feasibility, reliability or easiness of use, but few have used a specific acceptability measurement tool (Abboudi et al., 2013; Ahmed et al., 2011).

The aim of the present study was to assess the acceptability and usability of a procedural skill VR simulator using the Unified Theory of Acceptance and Use of Technology model (UTAUT) (Venkatesh et al., 2003; Venkatesh et al., 2012). The scenario for this VR simulator was designed to train scrub nurses in the preparation of the instrumentation table for a craniotomy. The goal was to predict its future use through

[☆] Authors' notes: The authors are particularly grateful to the engineers for their technical assistance and patience (A. Audinot, G. Claude, A. Lamercherie, J. Rojas-Balderrama) and the healthcare professionals of the neurosurgery department for their helpful comments (P.L. Henaux and B. Nogues).

* Corresponding author at: Université Rennes 2, Département de Psychologie, Place du Recteur Henri le Moal, 35043 Rennes Cedex, (France).

E-mail addresses: marie.stephanie.bracq@univ-rennes2.fr (M.-S. Bracq), estelle.michinov@univ-rennes2.fr (E. Michinov), bruno.arnaldi@irisa.fr (B. Arnaldi), benoit.caillaud@inria.fr (B. Caillaud), bernaud.gibaud@univ-rennes1.fr (B. Gibaud), valerie.gouranton@irisa.fr (V. Gouranton), pierre.jannin@univ-rennes1.fr (P. Jannin).

<https://doi.org/10.1016/j.nedt.2019.05.026>

Received 11 February 2019; Received in revised form 25 April 2019; Accepted 16 May 2019

0260-6917/ © 2019 Published by Elsevier Ltd.

motivation and behavioral intention, to anticipate obstacles to its adoption, and ultimately to optimize its design.

According to the UTAUT model, the main predictors of acceptance of a technology are Performance Expectancy (i.e. Perceived Utility), Effort Expectancy (i.e. Perceived Usability), Social Influence and Facilitating Conditions (i.e. available resources and support), Hedonic Motivation (i.e. fun or perceived pleasure), and Behavioral Intention, with three moderating factors: age, gender, and experience. The VR context adds a number of specific variables that may influence user experience in a VR environment (Williams et al., 2011). In the Components of User Experience model (CUE-model) (Thüring and Mahlke, 2007), the main elements are Perception of Instrumental Qualities, Perception of non-Instrumental Qualities and Emotional Reactions. In the present study, we assessed the non-instrumental qualities of the system as well as participants' reactions by assessing their subjective workload, their feeling of presence, their level of simulator sickness, and by asking them open-ended questions after the simulation session.

Based on the UTAUT model, behavioral intention to use the VR learning system was expected to be influenced by performance expectancy, effort expectancy, social influence, facilitating conditions and hedonic motivation, and to be moderated by age, gender and experience. According to this model, acceptability was expected to be higher for younger male participants, for scrub nurses and for VR users/video gamers. Moreover, according to the CUE-model, participants' intention to use the VR learning system was expected to be influenced by their level of subjective workload, feeling of presence and simulator sickness. Thus, acceptability was expected to be higher for participants with lower levels of subjective workload, stronger feelings of presence, and lower levels of simulator sickness.

2. Method

2.1. Participants

The VR simulator training system was first pre-tested with 5 VR experts (2 women and 3 men, mean age = 25.6), who were either PhD students in VR or VR engineers, familiar with best practices and interactions in virtual reality training systems. They tested the VR simulator and suggested design improvement. The simulator training system was then tested for acceptability by 16 non-expert users and 13 expert scrub nurses from the neurosurgery department of a French University Hospital.

2.2. Ethics

The study was approved by the Ethics Committee of the Hospital. Participants provided their informed consent for the study.

2.3. Apparatus

To ensure realism of the scenario, the formal methodology used for its design started with video observations recorded in the OR. The surgical activities shown on these videos were encoded using the *Surgery Workflow Toolbox* [Annotate] software¹ based on OntoSPM, a generic ontology for the domain of surgical processes (Gibaud et al., 2018). Scenarios resulting from these codings were then processed electronically using Test and Flip network (Caillaud, 2013) and #SEVEN (Claude et al., 2016), in order to model the variability of complex scenarios and the possible interactions between several users. A large screen allowed experimenters to monitor the execution of scenarios. The head-mounted display used for this study was the HTC Vive set.

¹ Surgery Workflow Toolbox* [Annotate] software: <https://b-com.com/en/bcom-surgery-workflow-toolbox-annotate>

2.4. Procedure

After reading and signing an information and consent form, participants completed a pre-experiment questionnaire to collect demographic data about gender, age, status, service and work experience, and to measure any prior interest or familiarity with technology, video games and virtual reality.

2.4.1. Instruction

Each simulation session began with a 3-min instruction video to explain VR concepts, use of controllers and metaphors. In order to be as instructive as possible, the explanations were illustrated with scenes from the VR pre-training scenario, as described below (see Fig. 1).

2.4.2. Pre-training session

After the instruction video, participants were presented with the VR pre-training scenario. Its goal was to enable participants to learn the interactions and to identify any contraindications. First, they assembled four Russian dolls inside each other. Then, they had to put three sugar cubes in a cup with sugar tongs, put the tongs back on a towel and fold it, cut a sheet of paper, and finally move a tray to the other end of the table. VR pre-training is important to limit mental load effects and stress that can result from an initial VR experience (Lackey et al., 2016). As it was assumed that most of the participants were novices in VR, this scenario was performed twice. Total time was not expected to exceed 20 min.

2.4.3. Test session of the instrumentation table scenario

After this initial scenario, participants were presented with the VR instrumentation table scenario (Fig. 2), lasting 20 min.

In accordance with the surgical protocol for a craniotomy, participants had to arrange the surgical instruments that were displayed in front of them or handed to them by the virtual circulating nurse. This scenario involved 138 surgical instruments and 50 different interactions between the user and the virtual environment.

All sessions were screen-recorded and each participant's performance was logged. After the simulation session, the participants completed a post-experiment questionnaire to measure their subjective workload, sense of presence, simulator sickness, and acceptability of the VR training system. They also took part in a recorded semi-structured interview (Fig. 3). The total duration of the experiment was 1 h30.

2.5. Measures

2.5.1. Personal innovativeness

In the pre-experiment questionnaire, participants were asked to complete the personal innovativeness scale (Agarwal and Karahanna, 2000). The four items of the scale were adapted to the present study, replacing "peers" by "colleagues" in the third item ("If I heard about a new IT (information technology) application, I would look for ways to try it out"; "In general, I am hesitant to try out new IT applications"; "Among my colleagues, I am usually the first to try new information technologies"; and "I like experimenting with new information technologies"). A 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*) was used. The reliability of the personal innovativeness scale was satisfactory (Cronbach's $\alpha = 0.79$) after removal of the reverse item related to hesitation to use new technologies.

2.5.2. Familiarity with video games and virtual reality

Based on the UTAUT model, it was expected that people who regularly used controllers and/or virtual environments would not have the same attitude towards the VR training simulator as those who were unfamiliar with them. In the briefing phase, participants were asked whether they played video games and if they had already used virtual reality (Yes/No answers).



Fig. 1. Photo of the pre-training scenario.

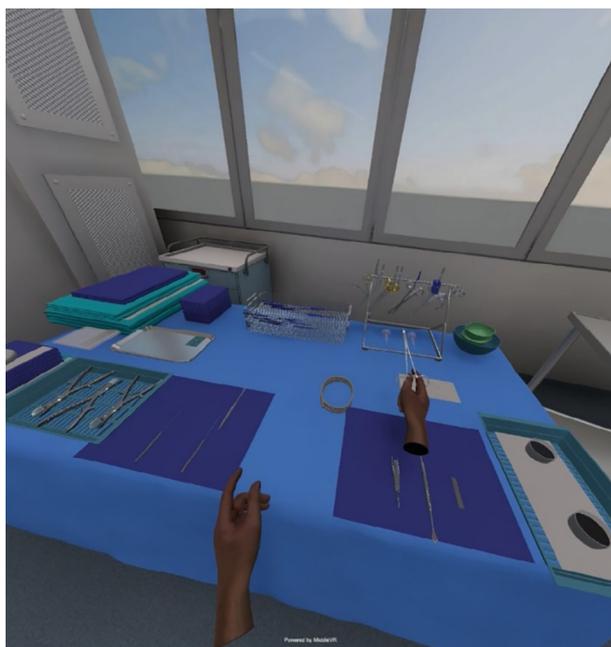


Fig. 2. Photo of the instrumentation table scenario.

2.5.3. Usability of the VR simulator

2.5.3.1. Subjective workload. The NASA Task Load Index (NASA-TLX; Hart and Staveland 1988; simplified version of Byers et al., 1989) was used to assess participants' subjective workload experience during the simulation session. The NASA-TLX scale consists of 6 items regarding workload (i.e. mental demand, physical demand, temporal demand, performance demand, effort, and frustration) rated on a 0–100 scale. The reliability of the NASA-TLX scale was acceptable (Cronbach's $\alpha = 0.65$). Removing the reversed item for performance demand would have improved it (Cronbach's $\alpha = 0.72$). Nevertheless, we preferred to keep the original scale to allow possible comparisons with other studies.

2.5.3.2. Feeling of presence. The feeling of presence was evaluated with

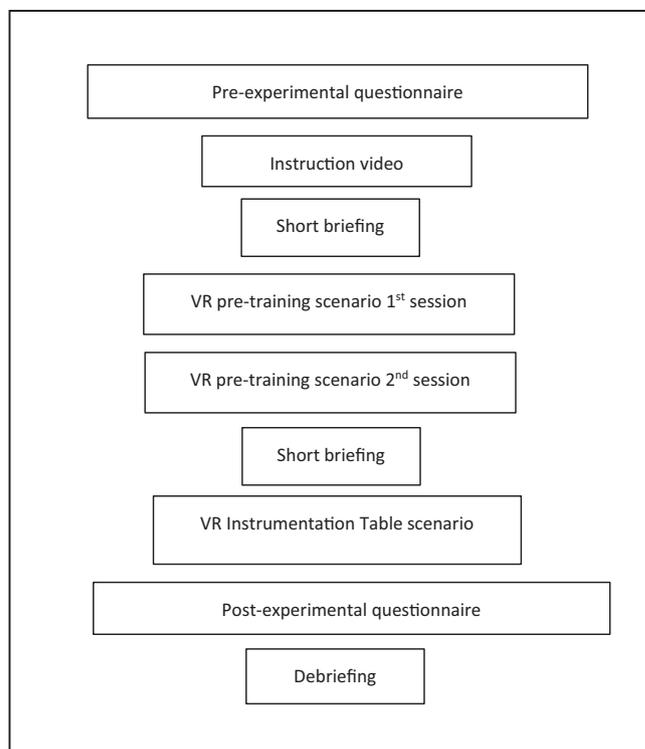


Fig. 3. Flowchart of the simulation session.

the SUS Questionnaire (Slater et al., 1994), which comprises 6 scales about the feeling of “being there”, the extent to which the virtual environment became the dominant reality, and the feeling of having visited a place rather than seeing images of it. For example: “I felt I was inside the virtual OR” from 1 (*Not at all, never*) to 7 (*Really a lot, a lot*), or “When I remember the experience, the virtual OR seems to be an image more than a place I’ve been to” from 1 (*Images I’ve seen*) to 7 (*A place I’ve been to*). The reliability of the presence scale was good (Cronbach's $\alpha = 0.79$).

2.5.3.3. Simulator sickness. Simulator sickness was assessed with 13 items from the Simulator Sickness Questionnaire (SSQ) (Bouchard et al., 2007; Kennedy et al., 1993) with a 4-point scale, from 0 (*no sign*) to 3 (*severe*). The original questionnaire contains 16 items but three of them, considered inappropriate for the scenario, were excluded: increased salivation, stomach awareness and burping. The oculomotor items were: fatigue, headache, eyestrain, difficulty focusing, difficulty concentrating, fullness of head, and blurred vision. The nausea items were: general discomfort, sweating, nausea, dizzy (eyes open), dizzy (eyes closed), and vertigo. The reliability of the global simulator sickness questionnaire was satisfactory (Cronbach's $\alpha = 0.71$) after deletion of two items related to “Sweating” and “Difficulty concentrating”. Responses to the 11 items were added together to form an overall simulator sickness score ranging from 0 to 33.

2.5.3.4. Task-completion times. Time needed for each participant to complete the task was isolated from logs as an objective index of usability of the VR simulator (Nielsen, 1993).

2.5.4. Acceptability of the VR simulator

2.5.4.1. Questionnaire of acceptability. The acceptability of the VR environment was measured with 21 items taken from the UTAUT2 questionnaire (Venkatesh et al., 2012), with a 7-point Likert scale from 1 (*Totally disagree*) to 7 (*Totally agree*). Price value and Habit were not assessed, as the study participants were not potential clients and

immersive VR is still considered a new technology, especially in a training context. Three dimensions of this questionnaire only concerned expert users: social influence (i.e. opinion of colleagues or hierarchy), facilitating conditions (i.e. resource and assistance provided by the hospital) and behavioral intention (i.e. recommendation to colleagues) (for examples of items, see Appendix 1). Reliability was assessed for the global UTAUT scale and for each of its dimensions. All Cronbach's alphas were satisfactory (see Table 2). For each dimension, items were aggregated to give an overall score.

2.5.4.2. *Interest for VR.* To measure participants' interest for VR, they responded to the same item before and just after the simulation session: "From my point of view, virtual reality can be useful for training", on a 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*).

2.5.4.3. *Participants' reactions.* Finally, participants were asked to list by order of importance three reasons why they would like to use the simulator again, and three reasons why they would not.

2.5.4.4. *Interviews.* Participants were finally invited to describe their experience with the VR system, discuss positive and negative aspects of the scenarios, and make suggestions to improve the training course.

3. Results

Statistical analyses were conducted with JASP (JASP Team, 2018). Descriptive statistics were computed, such as mean, maximum and minimum values, and standard deviations. We also compared independent subgroups using Mann-Whitney tests, pre- and post-simulation with paired Student's *t*-test, and correlations between main variables using Bravais-Pearson coefficients. Three non-expert users and two expert users were excluded due to technical problems for three of them, differences in health expertise for one, and because one participant was not a native French speaker.

3.1. Population description

Table 1 describes the sample of the present study. Non-expert users were 16 PhD students or engineers in informatics, robotics, psychology, ergonomics or business administration. They included eight women and eight men, with a mean age of 26.63 years. Seven of them had never used VR before, while nine had. Eleven of them had already played video games, while five never had. The mean value for personal innovativeness was 3.60/5.

The sample of expert users comprised 13 scrub nurses from the neurosurgery department of a French University Hospital. All of them

were women, with a mean age of 42 years. Four of them had already played video games, while nine never had. Similarly, four of them had already used VR, and nine never had. Their mean value for personal innovativeness was 3.38/5. No significant difference for personal innovativeness was observed between non-expert and expert users.

3.2. Subjective workload

Subjective workload for this VR training system was moderate for all dimensions, with some higher scores for mental demand, and lower scores for physical demand (see Fig. 4). No significant difference was observed between non-expert and expert users (see Table 1). For non-expert users, the mean value of total workload was 48.24/100 (*SD* = 14.22). For expert users, the mean value was 47.62/100 (*SD* = 11.12).

3.3. Feeling of presence

Feeling of presence in the VR simulator training system was globally good. No significant difference was observed between the two subgroups (see Table 2). For non-expert users, the mean value was 4.47/7 (*SD* = 1.14), and for expert users, it was 5.10/7 (*SD* = 0.96).

3.4. Simulator sickness

Simulator sickness was low in this VR training system, with the highest score for oculomotor fatigue. No difference was observed between non-expert and expert users (see Table 2). For non-expert users, the mean value for global simulator sickness was 3.94/33 (*SD* = 3.39), for nausea it was 0.56 (*SD* = 1.31), and for oculomotor problems it was 3.37 (*SD* = 2.70). For expert users, the mean value for global simulator sickness was 3.15/33 (*SD* = 2.97), for nausea it was 0.61 (*SD* = 1.12), and for oculomotor problems it was 2.54 (*SD* = 2.14).

3.5. Task-completion times

Mean task-completion time for the instrumentation table scenario was 13.44 min (*SD* = 3.65) for non-expert users, and 14.71 min (*SD* = 19.25) for expert users. No significant difference between subgroups was observed (Table 2).

3.6. Acceptability of the VR simulator

Acceptability of the VR simulator was rated good by both non-expert and expert users. Mean scores for each group are presented in Fig. 5. No significant difference was observed between the two

Table 1
Cronbach's alphas and comparison of mean values (Mann Whitney).

	Group 1 (n = 16)			Group 2 (n = 13)			Difference	
	Mean value	S.D.	Min – Max	Mean value	S.D.	Min–Max	U =	p =
Personal innovativeness	3.604	0.534	2.33–4.33	3.385	1.186	1–5	115.00	.642
Mental demand	59.47	22.84	22–87	67.23	16.77	32.5–94	89.00	.524
Physical demand	21.50	19.65	02–70	22.77	19.02	00–67.5	98.00	.809
Temporal demand	51.94	26.00	05–87	46.54	22.16	17.5–85	123.00	.417
Performance	53.66	18.75	25–90	45.96	17.06	12.5–72.5	123.00	.404
Effort	54.72	17.45	20–82.5	55.19	25.87	00–85	94.00	.996
Frustration	48.13	24.17	05–82.5	48.08	21.49	00–77.5	103.00	.982
Workload	48.24	14.22	29.3–68.3	47.62	11.12	24.2–61.7	109.00	.843
Presence	4.647	1.135	2.5–6.17	5.103	0.959	3.67–7.00	86.50	.455
Simulator sickness	3.94	3.29	0–11	3.15	2.97	0–10	118.00	.548
Performance expectancy	5.355	0.898	3.67–6.67	5.411	0.954	4–7	103.50	1.000
Effort expectancy	5.234	0.809	3.5–6.5	5.148	0.865	3–6.5	105.00	.982
Hedonic motivation	6.124	0.677	5–7	6.513	0.765	5–7	69.50	.120
Time	13.437	3.648	9.079–21.447	14.706	2.802	11–19.25	76.00	.232

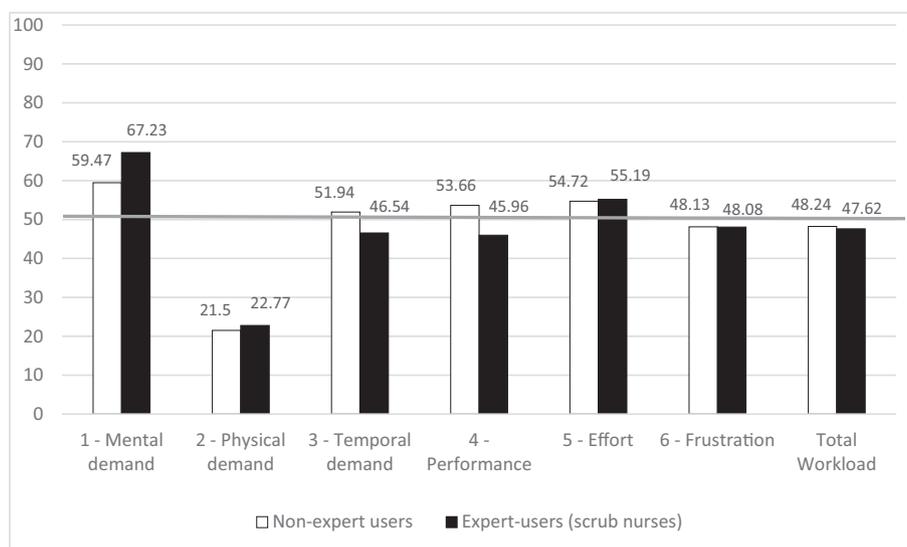


Fig. 4. Mean scores of NASA-TLX of non-expert and expert users.

subgroups (Table 1). For non-expert users, mean values were: 5.36/7 ($SD = 0.90$) for performance expectancy, 5.23/7 ($SD = 0.81$) for effort expectancy, and 6.12/7 ($SD = 0.68$) for hedonic motivation. For expert users, mean values were: 5.41/7 ($SD = 0.95$) for performance expectancy, 5.15/7 ($SD = 0.87$) for effort expectancy, 5.15/7 ($SD = 0.87$) for social influence, 5.87/7 ($SD = 0.62$) for facilitating conditions, 6.51/7 ($SD = 0.77$) for hedonic motivation, and 6.62/7 ($SD = 0.70$) for behavioral intention.

3.7. Interest in VR

Interest in VR for training was assessed pre- and post-simulation. The mean value was 4.48/5 ($SD = 0.57$) pre-simulation, and 4.79/5 ($SD = 0.34$) post-simulation. The difference is significant, $t(28) = -2.84, p = .008$, indicating that participants rated their interest in VR for training higher after the simulation session than before. No significant difference between non-expert and expert users was observed.

3.8. Participants' reactions

The occurrences of participants' reactions were summed (number of occurrences/total number of arguments) as follows.

For non-expert users, the major advantages of the VR system were “fun” (10/33), “pedagogical interest” (8/33) and “realism or immersion” (6/33). Other advantages concerned the discovery of a new technology, the originality of the tool, and a less expensive training system. The reported difficulties were “discomfort” (oculomotor fatigue) (10/24), “problem with interactions” (6/24), and “difficult to understand” (4/24). Other expressed limitations were feelings of disorientation and the repetitive nature of the tasks.

For expert users, the major advantages of the VR system were: “pedagogical interest” (11/28), “fun” (7/28), “innovation” (4/28), and “realism” (3/28). The other advantages noted were motivation to learn and the discovery of a professional environment under stress-free conditions. The difficulties noted were “discomfort” (oculomotor fatigue) (3/8), “lack of communication or discussion” (2/8), “some difficulties with gestures” (2/8).

3.9. Relationship between acceptability and user experience

Concerning the intention to use the VR simulator, behavioral intention was positively correlated with hedonic motivation and

facilitating conditions: the more participants enjoyed the environment, the more they wanted to use it again; the more participants thought they had the necessary resources or knowledge to use the environment, the more they wanted to use it again. Behavioral intention was negatively related to simulator sickness: the more participants felt simulator sick, the less they wanted to use the environment again.

Concerning the usability of the system, effort expectancy was negatively related to effort, frustration/anxiety and subjective workload: the less participants felt a sense of effort, the more they felt the environment was usable; the less participants felt frustrated or anxious, the more they felt the environment was usable; the less participants had a feeling of global workload, the more they felt the environment was usable; the more participants felt the simulation was physically demanding, the less they thought they had the necessary resources or knowledge to use it.

Concerning satisfaction with the system, hedonic motivation was positively correlated with mental and temporal demand: the more participants felt the simulation required attention or was mentally demanding, the more they enjoyed the environment; the more time pressure participants felt, the more they enjoyed the environment. Concerning user experience, the more participants felt immersed in the environment, the more they believed it was useful for their learning. Concerning simulator sickness, the more participants experienced simulator sickness, the less they thought they had the necessary resources or knowledge to use the environment. Simulator sickness was also positively correlated with performance: the more participants felt simulator sick, the more they thought they performed well.

Finally, some socio-demographic variables were related to task-completion time: the older participants took longer to complete the task, and women spent longer in the simulation session than men.

Table 2 shows the correlations between measures of acceptability and user experience for all participants.

3.10. Experience with the VR simulator

Comments provided by the participants after the simulation in the individual interviews demonstrate that the VR simulator was considered acceptable and usable. All participants perceived it as immersive and realistic: “It’s very immersive. You’re limited in space, but you want to explore, investigate. It’s a sign that I felt good”; “It represents the environment very well. I was impressed, everything is in the right place. It felt like being in an OR.” The verbatim record confirmed that the VR simulator was perceived as easy to use: “The interface is sometimes very

Table 2
Correlations for the studied variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1 Time	-																			
2 Mental Demand	0.358	-																		
3 Physical Demand	-0.109	0.007	-																	
4 Temporal Demand	0.087	0.653***	0.173	-																
5 Performance	-0.090	0.208	0.080	0.080	-															
6 Effort	0.414*	0.503**	0.028	0.648***	-0.098	-														
7 Frustration/anxiety	0.066	0.238	0.136	0.453*	0.008	0.366	-													
8 Workload	0.210	0.736***	0.373*	0.825***	0.257	0.711***	0.640***	-												
9 Presence	0.272	-0.200	-0.187	-0.345	-0.307	0.028	0.008	-0.272	-											
10 Performance Expectancy	0.290	0.044	-0.208	-0.279	0.185	0.095	-0.085	-0.083	0.497**	-										
11 Effort Expectancy	-0.236	-0.089	0.054	-0.337	-0.095	-0.417*	-0.432*	-0.383*	0.146	0.073	-									
12 Social Influence	-0.078	0.288	-0.220	-0.064	-0.114	0.076	-0.322	-0.114	0.464	0.442	0.568*	-								
13 Facilitating Conditions	-0.360	0.267	-0.581*	0.054	-0.165	-0.239	-0.104	-0.248	0.083	0.253	0.290	0.600*	-							
14 Hedonic Motivation	-0.025	0.399*	-0.271	0.412*	-0.100	0.150	0.186	0.242	-0.115	0.062	-0.061	0.179	0.643*	-						
15 Behavioral Intention	-0.164	0.260	-0.241	0.517	-0.319	0.092	0.364	0.241	-0.156	-0.103	-0.150	0.196	0.669*	0.860***	-					
16 Age	0.411*	0.285	0.168	0.110	-0.316	0.300	-0.123	0.125	0.068	-0.055	-0.054	-0.221	-0.526	0.187	-0.200	-				
17 Gender ^a	0.481**	0.187	-0.208	-0.016	-0.032	0.115	0.130	0.056	0.280	0.061	-0.281	NaN	NaN	0.256	NaN	0.467*	-			
18 Group ^b	0.195	0.193	0.034	-0.113	-0.215	0.011	-0.001	-0.024	0.217	0.031	-0.053	NaN	NaN	0.269	NaN	0.721***	0.556**	-		
19 Simulator Sickness	0.218	0.035	0.270	-0.005	0.378*	0.076	0.268	0.266	0.023	0.201	-0.111	-0.323	-0.806***	-0.329	-0.633*	-0.190	-0.131	-0.125	-	

Note. N = 29.

^a Gender: male = 1 and female = 2.

^b Group: non-expert user = 1 and expert user (scrub nurses) = 2.

* p < .05.

** p < .01.

*** p < .001.

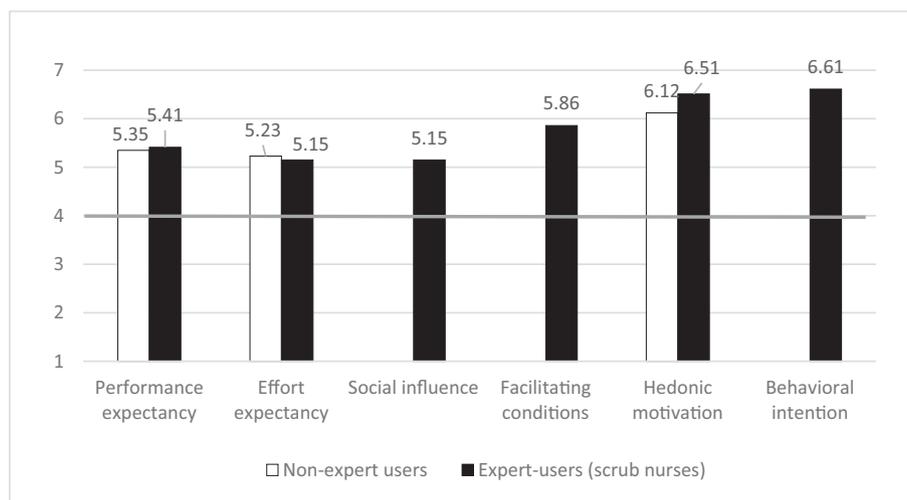


Fig. 5. Mean scores of each dimension of UTAUT of non-expert and expert users.

intuitive”, “It was the first time, I don’t play video games, but it’s easy to use”.

A number of participants also noted that it was an enjoyable experience, and that the VR system was fun: “You have to familiarize yourself with the system and you have to understand, but it doesn’t take long and it becomes fun”, “we do it while having fun, it’s fun and enjoyable”.

Concerning the difficulty using the VR system, participants said that it required considerable concentration: “It really looks like what we do, but it needed more thinking and concentration”, “It’s a completely different world, we have to read at the same time, it involves many abilities at the same time”. Some participants expressed some visual discomfort: “The helmet flickers a little, and it’s uncomfortable. It’s a bit heavy, you can feel the discomfort when you take it off”, “I was bothered by blurry vision”. Only one participant reported getting simulator sick: “Now I have a bit of a headache, and I feel queasy”.

Participants expressed an interest in the VR system for training in their professional practice: “For the younger ones, or for us in other specialties, it might help us prepare a table”; “For other acts, there would be much to learn from VR. That’s very interesting”. The scenario was constrained and a high level of guidance was provided for this simulation. Gradual removal of the latter for a progressive training was suggested, though it did not prevent participants from reflecting on the experience and making connections with their daily lives: “I was not bothered by the constraints of the scenario; in everyday life we constantly have to adapt”. However, a non-expert participant expressed doubts about the learning achieved and its transfer to the OR: “I’ve learned to do things in the virtual environment but I wouldn’t know how to do them in real life in the OR”.

4. Discussion

The objective of the present study was to assess the acceptability and usability of a VR simulator for procedural training. Based on the UTAUT model, we examined which dimensions were related to behavioral intention to use the VR simulator, and whether there were any differences between groups of expert and non-expert users.

Unlike the UTAUT model, age, gender and experience had no effect on the evaluation of acceptability of the VR simulator. This is congruent with other studies that also found that socio-demographic variables have no effect (Marchewka and Kostiwa, 2007; Sumak et al., 2010). For all dimensions and for both groups, mean values were higher than the scale’s midpoint value, showing acceptability of the VR simulator. This indicates that the scenario can be used to teach surgical procedures to a wide range of trainees. This is reinforced by the fact that there were no significant differences between expert and non-expert users in terms of subjective workload. Furthermore, expert users who were familiar with the OR and the meaning of the scenario felt as immersed as non-experts

who discovered it for the first time, showing that the scenario can be used for both initial and vocational training.

Many participants said that visual discomfort was the main drawback of the system; although simulator sickness was assessed as low, as observed in other studies (Bouchard et al., 2007), the oculomotor dimension was more perceptible. Although technical developments will lead to improvements in the quality of head-mounted devices (HMD) in the near future, simulator sickness remains an issue. The challenges and difficulties of using the VR system expressed in the final interviews show that this can have a negative effect on behavioral intention and on the perception of facilitating conditions.

Task-completion time is an objective measure of usability, and no significant differences were found between groups. Regardless of age or gender, all participants managed to learn how to use the simulator and successfully completed the simulation session.

Most participants stressed the pedagogical interest, fun and realism of the VR simulator, although non-expert users expressed some difficulty understanding the system, and expert users regretted a lack of communication, especially with the circulating nurse.

Another important result of the present study is that interest in VR for training increased between pre- and post-simulation. All participants rated VR for training higher after the simulation session than before, showing that the experience convinced them of the interest of developing this type of tool for training.

However, the purpose of the VR simulator and the scenario is not to replace other ways of training students or to be used on their own (Kneebone et al., 2007). In order to be fully effective, they have to be integrated into the curriculum, with pre-briefing and de-briefing depending on the specific educational objectives. As suggested by the participants, the scenario can be used by novices before their first OR experience (Breedt and Labuschagne, 2019), and the level of guidance can be adjusted as students progress.

Even though acceptability studies do not require large samples, one limitation of this study is its small sample size. Our expert group was composed of all the scrub nurses in one surgical department, limiting our sample to neurosurgery. It would be interesting in the future to conduct comparative studies with other medical specialties. Objective measures other than task-completion time could have been collected, such as hesitations, mistakes, or changes logged by the simulator, but the scenario was too constrained to allow any exploitation of these data. Qualitative data could also have been gathered using the “think aloud” technique. This option was considered but ultimately rejected for fear of significantly increasing participants’ subjective workload. However, immediate reactions were collected in post-simulation interviews and have been included in the present discussion.

5. Conclusion

This study demonstrates the acceptability of a VR simulator for procedural training skills. Unlike the UTAUT model, there was no effect of age, gender or expertise. Workload, immersion and simulator sickness were rated equally by all participants. This shows that our VR simulator, designed to teach surgical procedures, can be used extensively in initial or vocational training.

The next step will be to show whether this approach significantly improves procedural skills. Another line of development will be to design new scenarios to teach non-technical skills, such as situational awareness and communication in the OR, as suggested by some scrub

nurses in the post-test interviews.

Funding

This work is supported by the French Ministry of Research and Education as part of the “Laboratoires d’Excellence”, Labex CominLabs, SunSet project: “Scrub Nurse Non-Technical Skill Training System”.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Appendix 1. UTAUT questionnaire: assessed dimensions and examples of items

Dimension	Example of items
<i>Performance expectancy</i>	This learning environment allows you to learn more quickly This environment makes learning more effective
<i>Effort expectancy</i>	My interactions with this learning environment are clear and understandable This learning environment is easy to use
<i>Social influence</i>	My superiors think I should train in this learning environment My colleagues will support this learning environment
<i>Facilitating conditions</i>	I have the necessary knowledge to learn with this learning environment This learning environment seems compatible with the training methods used at hospital
<i>Hedonic motivation</i>	Using this learning environment is fun I enjoyed learning in this new environment
<i>Behavioral intention</i>	I am ready to use this learning environment for another training course If I had the opportunity, I would recommend this learning environment to my colleagues

References

Abboudi, H., Khan, M.S., Aboumarzouk, O., Guru, K.A., Challacombe, B., Dasgupta, P., Ahmed, K., 2013. Current status of validation for robotic surgery simulators - a systematic review: current status of validation for robotic surgery simulators. *BJU Int.* 111 (2), 194–205. <https://doi.org/10.1111/j.1464-410X.2012.11270.x>.

Agarwal, R., Karahanna, E., 2000. Time flies when you’re having fun: cognitive absorption and beliefs about information technology usage. *MIS Quarterly* 24 (4), 665–694. <https://doi.org/10.2307/3250951>.

Agha, R.A., Fowler, A.J., 2015. The role and validity of surgical simulation. *Int. Surg.* 100 (2), 350–357. <https://doi.org/10.9738/INTSURG-D-14-00004.1>.

Ahmed, K., Jawad, M., Abboudi, M., Gavazzi, A., Darzi, A., Athanasiou, T., Dasgupta, P., 2011. Effectiveness of procedural simulation in urology: a systematic review. *J. Urol.* 186 (1), 26–34. <https://doi.org/10.1016/j.juro.2011.02.2684>.

Bouchard, S., Robillard, G., Renaud, P., 2007. Revising the factor structure of the simulator sickness questionnaire. *Annu. Rev. Cyberther. Telemed.* 5, 117–122.

Breedt, S., Labuschagne, M.J., 2019. Preparation of nursing students for operating room exposure: a south African perspective. *Afr. J. Health Prof. Educ.* 11 (1), 22–26.

Byers, J.C., Bittner Jr., A.C., Hill, S.G., 1989. Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? In: Mital, A. (Ed.), *Advances in Industrial Ergonomics and Safety*. Taylor & Francis, London, pp. 481–485.

Caillaud, B., 2013. Surgical process mining with test and flip net synthesis. In: *Application of Region Theory (ART)*, pp. 43–54 (Barcelona, Spain).

Cipresso, P., Giglioli, I.A.C., Raya, M.A., Riva, G., 2018. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Front. Psychol.* 9. <https://doi.org/10.3389/fpsyg.2018.02086>.

Claude, G., Gouranton, V., Caillaud, B., Gibaud, B., Arnaldi, B., et al., 2016. Synthesis and simulation of surgical process models, studies in health technology and informatics. *IOS Press* 220, 63–70. <https://doi.org/10.3233/978-1-61499-625-5-63>.

Gibaud, B., Forestier, G., Feldmann, C., Ferrigno, G., Gonçalves, P., Haidegger, T., ... Jannin, P., 2018. Toward a standard ontology of surgical process models. *Int. J. Comput. Assist. Radiol. Surg.* 13 (9), 1397–1408. <https://doi.org/10.1007/s11548-018-1824-5>.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), *Advances in psychology*, 52. Human mental workload, pp. 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9).

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3 (3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3.

Kneebone, R.L., Nestel, D., Vincent, C., Darzi, A., 2007. Complexity, risk and simulation in learning procedural skills. *Med. Educ.* 41 (8), 808–814. <https://doi.org/10.1111/j.1365-2923.2007.02799.x>.

Lackey, S.J., Salcedo, J.N., Szalma, J.L., Hancock, P.A., 2016. The stress and workload of virtual reality training: the effects of presence, immersion and flow. *Ergonomics* 59 (8), 1060–1072. <https://doi.org/10.1080/00140139.2015.1122234>.

Marchewka, J.T., Kostiwa, K., 2007. An Application of the UTAUT Model for Understanding Student Perceptions Using Course Management Software. 7 (2), 13.

Nielsen, J., 1993. *Usability Engineering*. Academic Press, San Diego.

Slater, M., Usoh, M., Steed, A., 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3 (2), 130–144.

Sumak, B., Polancic, G., Hericko, M., 2010. An Empirical Study of Virtual Learning Environment Adoption Using UTAUT. pp. 17–22. <https://doi.org/10.1109/eLmL.2010.11>. February.

Team, J.A.S.P., 2018. *JASP (Version 0.9)[Computer Software]*. Retrieved from. <https://jasp-stats.org/>.

Thüring, M., Mahlke, S., 2007. Usability, aesthetics and emotions in human–technology interaction. *Int. J. Psychol.* 42 (4), 253–264. <https://doi.org/10.1080/00207590701396674>.

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: toward a unified view. *MIS Quarterly* 27 (3), 425–478. <https://doi.org/10.2307/30036540>.

Venkatesh, V., Thong, J.Y., Xu, X., 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly* 36 (1), 157–178.

Williams, M.D., Rana, N.P., Dwivedi, Y.K., Lal, B., 2011. Is UTAUT really used or just cited for the sake of it? A systematic review of citations of UTAUT’s originating article. In: ECIS, Retrieved from. <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1230&context=ecis2011>.