Full citation: Frederiksen JG, Sørensen SMD, Konge L, Svendsen MBS, Nobel-Jørgensen M, Bjerrum F, Andersen SAW. Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial. Surg Endosc. 2020 Mar;34(3):1244-1252.
DOI: 10.1007/s00464-019-06887-8

Title

Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery – a randomized trial

Running head

Training using immersive virtual reality

Authors

Joakim Grant Frederiksen, BSc.[a, b], Stine Maya Dreier Sørensen, MD [a], Lars Konge, PhD, [a, b], Morten Bo Søndergaard Svendsen, PhD, [a], Morten Nobel-Jørgensen, PhD. [c], Flemming Bjerrum, PhD, [a, d], Steven Arild Wuyts Andersen, PhD. [a, e],

[a] Copenhagen Academy for Medical Education and Simulation (CAMES), the Capital Region of Denmark, Copenhagen, Denmark.

[b] University of Copenhagen, Copenhagen, Denmark.

[c] Center for Computer Games Research, IT University of Copenhagen, Denmark.

[d] Dept. of Surgery, Herlev-Gentofte Hospital, Copenhagen, Capital Region, Denmark.

[e] Dept. of Otorhinolaryngology-Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark.

Corresponding author

Joakim Grant Frederiksen

Address: CAMES, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark.

Telephone number: +45 22 58 86 88

Email: Joakim.grant.f@gmail.com

Abstract

Background: Virtual reality simulators combined with head mounted displays enable highly immersive virtual reality (IVR) for surgical skills training, potentially bridging the gap between the simulation environment and real-life operating room conditions. However, the increased complexity of the learning situation in IVR could potentially induce high cognitive load (CL) thereby inhibiting performance and learning. This study aims to compare CL and performance in IVR and conventional virtual reality (CVR) simulation training.

Methods: A randomized controlled trial of residents (n=31) performing laparoscopic salpingectomies with an ectopic pregnancy in either IVR or CVR simulation. CL was estimated by secondary-task reaction time at baseline, and during non-stressor and stressor phases of the procedure. Simulator metrics were used to evaluate performance.

Results: CL was increased by 66 % and 58 % during IVR and CVR simulation, respectively ($p<0.001$), compared to baseline. A light stressor induced a further increase in CL by 15.2 % and a severe stressor by 43.1 % in the IVR group compared to 23 % (severe stressor) in the CVR group. IVR also caused a significantly worse performance on most simulator metrics.

Conclusion: IVR simulation training induces a higher CL and results in a poorer performance than CVR simulation training in laparoscopy. High extraneous load and element interactivity in the IVR are suggested as mechanisms explaining this finding. However, IVR offers some potential advantages over CVR such as more real-life conditions but we only recommend introducing IVR in surgical skills training after initial training in CVR.

Keywords: Laparoscopic surgical skills training; Immersive virtual reality; Cognitive load; Head mounted device; Simulation

Introduction

The traditional model of surgical training, the apprenticeship model, may be inadequate for complex procedures like those performed with laparoscopic surgery[1] and is also associated with increased error and risk of complications[2-4]. Simulation-based training, like virtual reality (VR) simulator training, enables trainees to learn surgical skills in a risk-free environment and increases technical proficiency, improves surgical trainees' operating performance, and decreases operating time in laparoscopic procedures[5-7]. As a result, simulation-based training is increasingly being implemented into surgical training programs[8].

A challenge for VR simulation-based training is the substantial gap between the simulation-environment and the real-life operating room (OR)[9]. Technology such as head-mounted displays (HMDs) in combination with VR laparoscopic simulators can be used to create immersive VR (IVR) simulation where the surroundings and environment from a real OR are integrated into the VR learning experience[10]. This can potentially bridge the gap between simulation and real-life OR experience, creating a more realistic and complete training environment. However, as the realism of the simulation increases, the strain on the trainees' cognitive capacity might increase as well, challenging actual learning.

To investigate learning in an IVR environment we used cognitive load (CL) theory as a framework. Most literature on CL theory considers three different types of CL[11]: The *intrinsic* load of the learning task (inherent to the complexity of the task), the *extraneous* load of the learning situation (a result of superfluous processes that do not directly contribute to learning), and the *germane* load of the learning process (mental schema formation for the actual learning)[11]. IVR skill training inherently has a high level of element interactivity[11], as for example staff and instruments can be (perceived as) possible elements to interact with, contributing to a high intrinsic CL. Furthermore, distractions in the immersive environment add extraneous CL. This possible increase in intrinsic and extraneous load during training in IVR compared with CVR could result in cognitive overload, inhibiting learning, due to there being no cognitive capacities allocated to germane load[11-15]. Other studies have investigated CL in less immersive surgical simulation training, but none have investigated CL in IVR nor explored the consequences for performance. Moreover, as motion sickness has been associated with IVR[16] this was investigated as well.

The applications for IVR training are many and could possibly contribute to surgical training in a number of ways. In this study, we wanted to investigate CL and performance in a randomized controlled study during IVR and CVR training of laparoscopic salpingectomy for ectopic pregnancy to investigate the use of IVR in novice training.

Material and methods

Study Design

A single-center randomized trial was designed according to the CONSORT statement (Figure 1)[17].

The trial was exempted for ethical approval by the Regional Biomedical Research Ethics Committee (H-17041390) and was registered at clinicaltrials.gov (NCT03721094).

Participants

Participants were first-year residents without previous laparoscopic experience (Table 1). They were recruited by invitations sent to their clinical departments as well as junior doctors' surgical interest groups. Participation was voluntary and not financially compensated. Participants received written and verbal information prior to giving informed consent for study enrollment.

The inclusion criterion was that the participant was a first-year resident. The exclusion criteria were: 1) Previous participation in any projects involving laparoscopic training, 2) Prior experience with laparoscopic surgery (having performed one or more laparoscopic procedures as primary surgeon, including supervised procedures), 3) Not speaking Danish on a conversational level. Upon inclusion, all participants were assigned a personal trial identification number before randomization.

The study took place at the Simulation Center at Rigshospitalet, Copenhagen Academy for Medical Education and Simulation (CAMES) during February through June 2018[18].

Interventions

Participants first completed a demographics questionnaire (Table 1). Secondly, they received hands-on introduction to the CVR laparoscopic simulator and trained four basic skill modules (touching, grasping, cutting, and fine dissection). These tasks were each preceded by a video instruction and each task was practiced repeatedly for 12 minutes supervised by the main investigator (JGF), who could resolve any technical issues the participants could experience with the simulator. Finally, a video instruction on the ectopic salpingectomy procedure was reviewed followed by one supervised hands-on procedure in the simulator (conventional VR setup).

Next, participants were randomized to the intervention or the control group and completed three attempts of the same procedure (laparoscopic salpingectomy due to an ectopic pregnancy) without feedback or guidance. The intervention group performed the procedures while wearing a HMD for the immersive VR environment (IVR group), whereas the control group performed the procedures with conventional VR (CVR group) (Figure 2A, Figure 2B). We chose the laparoscopic salpingectomy procedure because it could be standardized and has a bleeding complication that could be controlled and triggered in relation to the scenarios. Finally, the procedure was relevant and sufficiently challenging for the study population.

In the IVR environment, four different 360-degrees videos were in sequence played as backdrop during the procedure. The videos reflected real life situations in the operating room with two videos representing calm periods, one video representing a light stressor and one video representing a severe stressor with a bleeding (2 ml/s) being triggered in the simulation (video descriptions, Table 2). None of the videos required direct involvement or response by the participant. All videos were looped with actors starting and ending in the same positions so they could be played in sequence without noticeable shifts (sequence of events, Figure 3).

In the CVR setup, the participant practices in the regular and calm environment of the simulation center and at the same time during the procedure, the bleeding (2 ml/s) is triggered as a stressor (Figure 3).

Material and equipment

The conventional VR laparoscopic simulation setup consisted of 2 Simball 4D joysticks (G-coder Systems, Gothenburg, Sweden) connected to a 27” monitor with a computer running the LapSim® software version 2016. (Surgical Science, Gothenburg, Sweden). The module used was the laparoscopic salpingectomy due to an ectopic pregnancy. A second computer running the

TeamSim® (Surgical Science, Gothenburg, Sweden) VR software version 2016 allowed for bleeding during the procedure to be initiated by the investigator.

The immersive VR laparoscopic simulation setup consisted of the above with an additional computer handling the Oculus Rift® (Oculus VR, Irvine, USA) and playback of 360-degrees videos of the operating room in Unity 3D (Unity Technologies, San Francisco, USA). In the immersive VR environment, the visual output (laparoscopic view) from the simulator was projected as an overlay on the tower in the OR.*

The 360-degrees videos were filmed at an operating room at the Dept. of Abdominal Surgery, Rigshospitalet, with a 360-degree camera (Thor, Absolute Zero, Copenhagen, Denmark). Five different video sequences were recorded at 60 fps in 7680x3840 resolution.

Outcomes

The primary outcome was CL estimated by secondary task reaction time[19]. We chose to use secondary task reaction time as a measurement for CL, as studies have shown this to be a reliable method for measuring CL in surgical skills training[20-22], in contrast to subjective methods such as questionnaires that has limited use in measuring changes to CL during a procedure because they are administered at the end of a procedure[23]. We used an external and commercially available reaction timer (American Educational Products LLC, USA) to measure participants response time (in hundredth seconds) to an auditory stimulus (a beep). The participants responded to the auditory stimulus by pressing a pedal next to the foot pedal used for cauterization. Reaction time was measured before and after the simulation to provide an individual baseline, and during the

* Videos as online content: https://youtu.be/-PJKZJz6cc0 (calm phase), https://youtu.be/COtp3x0MWoI (light stressor), https://youtu.be/LqNV8euKTxM (severe stressor), https://youtu.be/TEAdtLBbixE (full procedure)

simulation at t=80 s, t=130 s, t=180 s and t=240 s (representing two calm phases and two phases with stressors during the IVR/three calm phases and one stressor phase during the CVR, see Figure 3). All reaction time measurements were done in series of four repeated measurements. It was also noted if the reaction time was measured while the participant was using the foot pedal for cauterization in the simulation.

Performance data was collected in the form of simulator metrics. We focused on time to completion of procedure, damage to surrounding tissue (diathermy damage and blood loss), and efficiency of instrument movements (Table 4).

Finally, all participants completed a motion sickness questionnaire[24] after their final procedure.

Sample size calculation

There is no generally accepted model for sample size calculations for linear mixed models with repeated measurements. We therefore based the sample size on data from previous studies[25, 26], estimating that 15 participants would be needed in each group to significantly detect a 2 % change in reaction time relative to individual baseline (the primary outcome).

Randomization

Participants were randomized 1:1 to the intervention (IVR) or control (CVR) group using a web-based service Sealed Envelope[tm] (Sealed Envelope™, London, UK). The allocation sequence was computer-generated and used variable block sizes of two and four, a sequence that was kept secret throughout the trial. Randomization was stratified for sex, as previous studies have demonstrated this may impact initial laparoscopic simulator performance[27].

Blinding

Participants and the principal investigators were blinded to performance metrics by the simulator during the supervised procedure and the three test procedures. Participants and the data collector could not be blinded to the allocation/intervention. The investigator in charge of the statistical analysis (SA) was blinded to participants' group allocation.

Statistical methods

The data was analyzed using SPSS® version 23.0 (IBM, Armonk, NY, USA) for Windows/Mac. Reaction time measurements during simulation were calculated relative to the individual baseline. Linear mixed models (LMM) were used in the analysis of reaction time (CL) and simulator metrics due to repeated measurements. Models were iteratively built, investigating different factors and interactions. For the CL analysis, the final model included group (intervention/control), procedure number (1–3), time of measurement (t=80/130/180/240 s), measurement number (1–4), measurement while cauterizing (yes/no), and group * time interaction. For the performance analysis, the final model included group, procedure number, and group * procedure number. For the motion sickness questionnaire data, independent samples t-test was used.

Estimated marginal means and p-values of the LMM are reported. For pairwise comparison within the individual factors in the LMM models, least significant difference p-values are reported. P-values <0.05 were considered statistically significant.

Results

Participant demographics are presented in Table 1. All 31 participants recruited for study completed the introduction and the three procedures.

Cognitive load

Overall, we found a significant difference in CL between the IVR and CVR groups, with IVR inducing a CL 7.9 % higher than CVR, $p<0.001$ (Table 3). For both groups, repeated practice caused an average of 6.4 % decrease in CL per repetition. CL measured during cauterization was 6.1 % further increased (Table 3).

In both groups, stressors significantly increased CL (Figure 4): in the IVR group, the light stressor (background conversation) further increased CL by 15.2 % and the severe stressor further increased CL by 43.1 % thereby almost doubling reaction time compared to non-simulation baseline; in the CVR group, the stressor further increased CL by 23.0 %, which was significantly less than the increase found in the IVR group, $p<0.001$. For both groups, there were no statistically significant differences in CL during the different non-stressor periods. We chose to use reaction time relative to baseline rather than the actual reaction time, as participants' reaction time varies from day to day based on tiredness, mood, and other factors. This is also reflected in the mean and standard deviation of the baseline measurements (39.2 sec, SD 5.8) and measurements during the procedure (62.9, SD 16.5).

Performance

For all simulator metrics, the CVR group significantly outperformed the IVR group (Table 4): the IVR group spent more time before completion of the procedure, induced more damage to the tissue, had a higher blood loss, and were less efficient with their hand movements. In both groups, repeated practice increased performance for total time ($p<0.001$) and blood loss ($p<0.001$). None of the other simulation metrics changed significantly during the three procedures.

Motion sickness

For the combined motion sickness score, there were no significant difference between the two groups (conventional VR: 24.1 points; immersive VR 25.7 points. p=0.62). No significant differences were found for the four sub-scores (gastrointestinal, central, peripheral and sopite-related).

Discussion

Cognitive load

Our study demonstrates that immersive VR induces a significantly higher cognitive load than conventional VR in general and especially during stressor phases. This can according to CL theory most likely be attributed to a perceived higher possible number of elements to interact with, such as staff and instruments (causing intrinsic load), as well as distractions in the immersive environment (causing extraneous load). During the severe stressor (triggering of bleeding), there was a substantial increase in CL in both IVR and CVR, which can be explained by the similar increase in task-complexity (higher intrinsic load). However, the further increase in CL observed in the IVR group is likely a result of the distraction by the immersive environment, causing increased extraneous load. This corroborates finding from other studies and suggests that highly complex learning environments induce a high level of CL[25]. Considering that the typical target group for simulation-based surgical skills training are novices and that novices are especially at risk of cognitive overload[28], a high level of CL might be undesirable during initial training. Altogether, this points towards using conventional VR simulation initially because it induces less CL and introducing immersive VR secondly to further bridge the gap between simulation and real-life OR experience. Finally, we also found that repeated practice decreased CL, in line with other reports[26, 29]. Interestingly, CL decreased at the same rate for both IVR and CVR, which could indicate that IVR has the potential to prepare trainees for the complex environment in real-life OR.

Performance

Performance was found to be significantly worse in the IVR group compared with the CVR group, indicating that the high CL of the IVR simulation environment has a negative effect on performance. On the other hand, it could also be suggested that the novelty of training with HMDs and unfamiliarity with the technology could be an explanation. However, if this was the case, we would have expected a larger improvement in performance with repeated practice as participants familiarized with the new equipment compared with the CVR group, however we found the performance of the two groups to improve at the same rate. In both groups, repeated practice resulted in a better performance regarding completion time and blood loss. Efficiency of hand movements did not improve, which is most likely explained by the participants not having spare cognitive capacity to focus on improving instrument handling and mainly focusing on controlling the bleeding.

Motion sickness

We found no significant difference between IVR and CVR in relation to motion sickness. In fact, our IVR setup seemed not to induce any motion sickness which may be due to minimal head movements compared to IVR video games where motion sickness has been an issue. This corroborates the finding from the only other study using same type of setup[10].

Strengths and limitations

At current, the simulation setup we have used represents the latest high-end, commercially available HMD and laparoscopic VR simulator for providing immersive VR. Further, our setup innovatively combines the TeamSim software to correspond with playback of 360° video, allowing real-time changes in the simulation conditions to correspond to changes in the simulation environment such as instrument failure, stopping/starting bleedings etc. We think this increase immersion and realism, better preparing learners for real-life conditions and challenges, as their training resembles a real-life OR. One of the current limitations of the IVR setup is that it does not simulate (or allow), view of the participants' hands, physical instruments, or foot-pedal (Figure 2A). It did, however, not seem to be a practical issue as participants rarely shifted their view from the screen and quickly memorized the position of the foot-pedal.

Perspectives

In the medical education literature, HMDs for VR simulation are still considered novel and the only reported use has been for team training scenarios[30, 31]. Even though the use of advanced VR HMDs in higher education is predicted to take place in 2018-2019[32] it is already in use in other fields such as psychology[33], aviation[34], and military[35] and for reducing stress in preoperative patients[36]. In surgical procedural training, we are aware of two other studies on advanced VR HMDs: In 2016, Sankaranarayanan et al[31] explored their Gen2-VR© comparing it to conventional VR, and found that participants performed poorer in the IVR setup when performing a peg transfer task. It is worth noting that this study used a computer-generated virtual environment and distractions in the immersive environment consisted of music been played and instruments malfunctioning. In 2017, Huber et al[10] in a non-randomized study also reported poorer performance by participants in an IVR group compared with conventional VR when performing laparoscopic fine dissection and cholecystectomy. They used a background video depicting routine processes during a standard laparoscopic procedure but did not use further distractions in the IVR nor did the investigate CL as a mechanism for explaining the poorer performance of the IVR group.

Immersive VR is a new addition to the current range of simulation-based training options for surgical skills training, with the benefit of more realistic conditions because of elements such as OR surroundings and team interactions. The increased realism of immersive VR likely causes a higher sense of presence, which can potentially make VR training more effective[16]. On the other hand, as the complexity of the scenarios in IVR increases, so does the risk of cognitive overload. To combat this different instructional design strategies to lower CL could be explored[11, 37] and studies investigating CL during IVR in more simulator proficient trainees are needed. Moreover, IVR might better prepare surgical novices for the environment in the OR than CVR, but studies are required to confirm this.

Future perspectives of IVR could be interactive scenarios with feedback from staff and equipment in the simulation. Thus, actions and mistakes on the VR simulator would have consequences in the virtual environment, this in turn could lead to an even higher feeling of immersion.

Conclusion

Immersive VR simulation for laparoscopic surgical training induces a higher cognitive load and results in worse performance compared with conventional VR simulation training of laparoscopic novices. The immersive environment most likely causes both additional intrinsic and extraneous CL, which could be disadvantageous in the initial training of novices and should optimally be introduced after basic training with conventional VR simulation. However, immersive VR has the potential to prepare surgical trainees for real-life conditions and challenges, and future developments of the immersive VR technology will potentially include interactions with staff and equipment, further bridging the gap between simulation and real-life.

Acknowledgments

The authors would like to thank the following for their help in producing the 360º videos; Fanny Klara Karolina Kullberg, Mai Homburg and Christoffer Skov Olesen for their portrayal of the staff in the operating room, Helle Skovgaard and Lars Bo Svendsen for their help in precuring a location for the recording of the videos and Absolute Zero© for their help in recording and processing the videos.

Disclosures

Joakim Grant Frederiksen, Stine Maya Dreier Sørensen, Lars Konge, Morten Bo Søndergaard Svendsen, Morten Nobel-Jørgensen, Flemming Bjerrum and Steven Arild Wuyts Andersen have no conflicts of interest or financial ties to disclose

References

1. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, et al. (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training Ann Surg 241:364-72
2. Reichenbach DJ, Tackett AD, Harris J, Camacho D, Graviss EA, Dewan B, et al. (2006) Laparoscopic colon resection early in the learning curve: what is the appropriate setting? Ann Surg 243:730-5; discussion 5-7 doi:10.1097/01.sla.0000220039.26524.fa
3. Watson DI, Baigrie RJ, Jamieson GG (1996) A learning curve for laparoscopic fundoplication. Definable, avoidable, or a waste of time? Ann Surg 224:198-203
4. Kasotakis G, Lakha A, Sarkar B, Kunitake H, Kissane-Lee N, Dechert T, et al. (2014) Trainee participation is associated with adverse outcomes in emergency general surgery: an analysis of the National Surgical Quality Improvement Program database Ann Surg 260:483-90; discussion 90-3 doi:10.1097/sla.0000000000000889
5. Varas J, Mejía R, Riquelme A, Maluenda F, Buckel E, Salinas J, et al. (2012) Significant transfer of surgical skills obtained with an advanced laparoscopic training program to a laparoscopic jejunojejunostomy in a live porcine model: feasibility of learning advanced laparoscopy in a general surgery residency Surgical Endoscopy 26:3486-94 doi:10.1007/s00464-012-2391-4
6. Zendejas B, Cook DA, Bingener J, Huebner M, Dunn WF, Sarr MG, et al. (2011) Simulation-based mastery learning improves patient outcomes in laparoscopic inguinal hernia repair: a randomized controlled trial Ann Surg 254:502-9; discussion 9-11 doi:10.1097/SLA.0b013e31822c6994
7. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR (2013) Virtual reality training for surgical trainees in laparoscopic surgery Cochrane Database Syst Rev:CD006575 doi:10.1002/14651858.CD006575.pub3
8. Sirimanna P, Gladman MA (2017) Development of a proficiency-based virtual reality simulation training curriculum for laparoscopic appendicectomy ANZ J Surg 87:760-6 doi:10.1111/ans.14135
9. Moorthy K, Munz Y, Adams S, Pandey V, Darzi A (2005) A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre Ann Surg 242:631-9
10. Huber T, Paschold M, Hansen C, Wunderling T, Lang H, Kneist W (2017) New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff Surg Endosc 31:4472-7 doi:10.1007/s00464-017-5500-6
11. van Merrienboer JJ, Sweller J (2010) Cognitive load theory in health professional education: design principles and strategies Med Educ 44:85-93 doi:10.1111/j.1365-2923.2009.03498.x
12. Cowan N (2014) Working Memory Underpins Cognitive Development, Learning, and Education Educational psychology review 26:197-223 doi:10.1007/s10648-013-9246-y
13. Choi H-H, van Merriënboer JJG, Paas F (2014) Effects of the Physical Environment on Cognitive Load and Learning: Towards a New Model of Cognitive Load Educational Psychology Review 26:225-44 doi:10.1007/s10648-014-9262-6
14. Kulasegaram KM, Grierson LE, Norman GR (2013) The roles of deliberate practice and innate ability in developing expertise: evidence and implications Med Educ 47:979-89 doi:10.1111/medu.12260
15. de Jong T (2010) Cognitive load theory, educational research, and instructional design: some food for thought Instructional Science 38:105-34 doi:10.1007/s11251-009-9110-0
16. Dargar S, Kennedy R, Lai W, Arikatla V, De S (2015) Towards immersive virtual reality (iVR): a route to surgical expertise J Comput Surg 2 doi:10.1186/s40244-015-0015-8
17. Schulz KF, Altman DG, Moher D (2011) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials Int J Surg 9:672-7 doi:10.1016/j.ijsu.2011.09.004
18. Konge L, Ringsted C, Bjerrum F, Tolsgaard MG, Bitsch M, Sorensen JL, et al. (2015) The Simulation Centre at Rigshospitalet, Copenhagen, Denmark J Surg Educ 72:362-5 doi:10.1016/j.jsurg.2014.11.012
19. Brunken R, Plass JL, Leutner D (2003) Direct Measurement of Cognitive Load in Multimedia Learning Educational Psychologist 38:53-61 doi:10.1207/S15326985EP3801_7
20. Rojas D, Haji F, Shewaga R, Kapralos B, Dubrowski A (2014) The impact of secondary-task type on the sensitivity of reaction-time based measurement of cognitive load for novices learning surgical skills using simulation Studies in health technology and informatics 196:353-9

21. Naismith LM, Cavalcanti RB (2015) Validity of Cognitive Load Measures in Simulation-Based Training: A Systematic Review Acad Med 90:S24-35 doi:10.1097/ACM.0000000000000893
22. Haji FA, Rojas D, Childs R, de Ribaupierre S, Dubrowski A (2015) Measuring cognitive load: performance, mental effort and simulation task complexity Med Educ 49:815-27 doi:10.1111/medu.12773
23. Naismith LM, Cheung JJ, Ringsted C, Cavalcanti RB (2015) Limitations of subjective cognitive load measures in simulation-based procedural training Med Educ 49:805-14 doi:10.1111/medu.12732
24. Gianaros PJ, Muth ER, Mordkoff JT, Levine ME, Stern RM (2001) A questionnaire for the assessment of the multiple dimensions of motion sickness Aviation, space, and environmental medicine 72:115-9
25. Andersen SA, Mikkelsen PT, Konge L, Caye-Thomasen P, Sorensen MS (2016) Cognitive Load in Mastoidectomy Skills Training: Virtual Reality Simulation and Traditional Dissection Compared J Surg Educ 73:45-50 doi:10.1016/j.jsurg.2015.09.010
26. Andersen SAW, Konge L, Sorensen MS (2018) The effect of distributed virtual reality simulation training on cognitive load during subsequent dissection training Medical teacher:1-6 doi:10.1080/0142159x.2018.1465182
27. Ali A, Subhi Y, Ringsted C, Konge L (2015) Gender differences in the acquisition of surgical skills: a systematic review Surg Endosc 29:3065-73 doi:10.1007/s00464-015-4092-2
28. Haji FA, Cheung JJ, Woods N, Regehr G, de Ribaupierre S, Dubrowski A (2016) Thrive or overload? The effect of task complexity on novices' simulation-based learning Med Educ 50:955-68 doi:10.1111/medu.13086
29. Andersen SA, Mikkelsen PT, Konge L, Caye-Thomasen P, Sorensen MS (2016) Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation Laryngoscope 126:E74-9 doi:10.1002/lary.25449
30. Abelson JS, Silverman E, Banfelder J, Naides A, Costa R, Dakin G (2015) Virtual operating room for team training in surgery Am J Surg 210:585-90 doi:10.1016/j.amjsurg.2015.01.024
31. Sankaranarayanan G, Li B, Manser K, Jones SB, Jones DB, Schwaitzberg S, et al. (2016) Face and construct validation of a next generation virtual reality (Gen2-VR) surgical simulator Surg Endosc 30:979-85 doi:10.1007/s00464-015-4278-7
32. Johnson L, Adams Becker S, Cummins M, Estrada V, Freeman A, Hall C. NMC Horizon Report: 2016 Higher Education Edition. Austin, Texas: The New Media Consortium; 2016.
33. Nararro-Haro MV, Hoffman HG, Garcia-Palacios A, Sampaio M, Alhalabi W, Hall K, et al. (2016) The Use of Virtual Reality to Facilitate Mindfulness Skills Training in Dialectical Behavioral Therapy for Borderline Personality Disorder: A Case Study Frontiers in psychology 7:1573 doi:10.3389/fpsyg.2016.01573
34. Lofaro RJ SK (2009) Civil aviation: flight simulators and training. In: Vincenzi DA WJ, Mouloua M, Hancock PA (ed)^(eds) Human factors in simulation and training, vol., vol, edn. CRC Press, Boca Raton, FL, p^pp 257-86.
35. Schmorrow D ND, Lackey SJ, Allen RC, Norman K, Cohn J (2009) Virtual reality in the training environment. In: Vincenzi DA WJ, Mouloua M, Hancock PA (ed)^(eds) Human factors in simulation and training, vol., vol, edn. CRC Press, Boca Raton, FL, p^pp 201-28.
36. Bekelis K, Calnan D, Simmons N, MacKenzie TA, Kakoulides G (2017) Effect of an Immersive Preoperative Virtual Reality Experience on Patient Reported Outcomes: A Randomized Controlled Trial Ann Surg 265:1068-73 doi:10.1097/sla.0000000000002094
37. Sweller J (1999) Instructional design in technical areas vol., vol, edn. ACER Press, Camberwell, Vic.

Figure legends

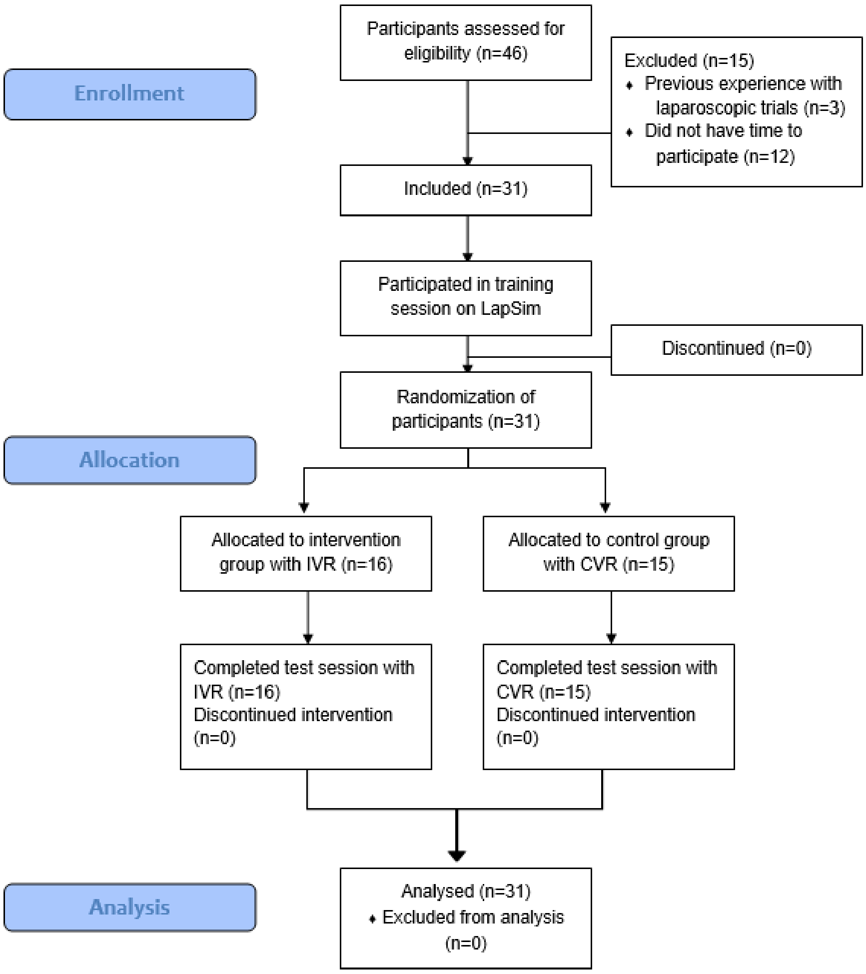

Figure 1: Trial flowchart in accordance with the CONSORT statement.

Figure 2A: External view of two participants training in an IVR and CVR session. Computers handling TeamSim and Oculus Rift is outside the picture. NOTE: no participants trained simultaneously. B: Participants personal 360º immersive view of the operating room during the IVR sessions.

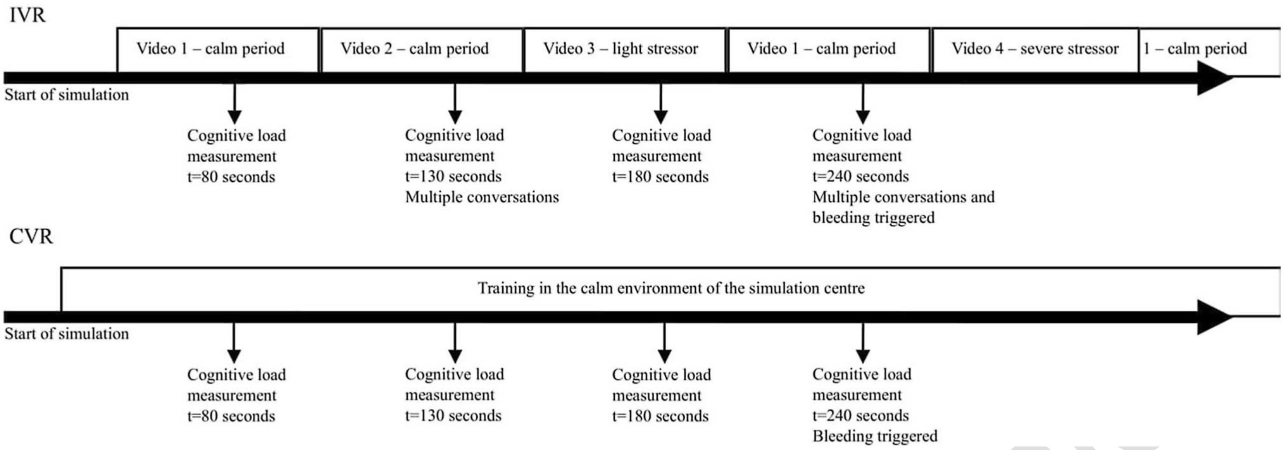

Figure 3: Sequence of videos and cognitive load measurements for the immersive virtual reality group (IVR) and conventional virtual reality group (CVR). After the last video 2, the sequence alternated between video 2 and 1 until completion of the procedure. For a description of the videos see Table 2.

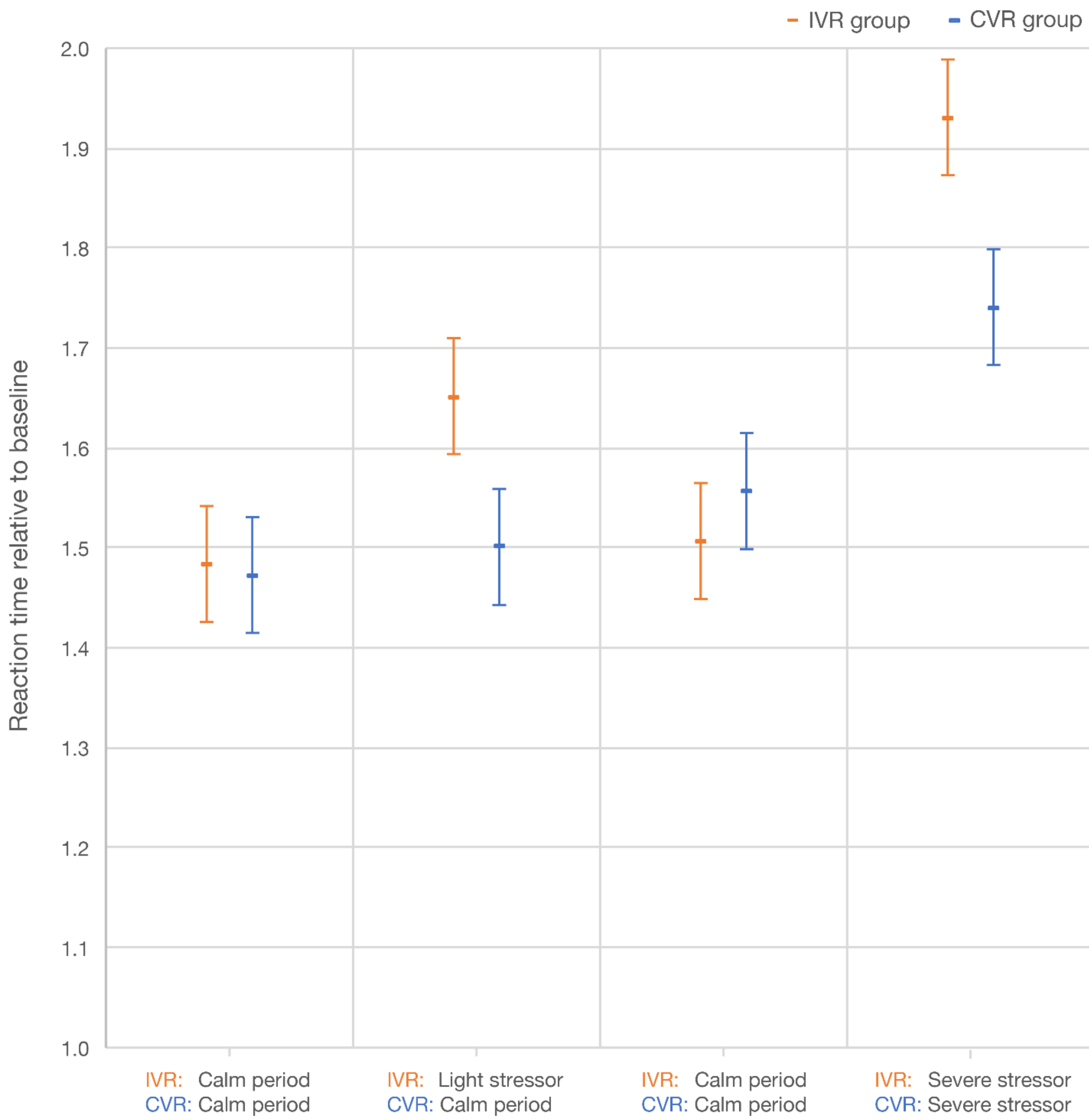

Figure 4: Means plot with 95% CI as error bars over reaction time relative to baseline measurements.

Table 1. Participant baseline characteristics

	Intervention group Immersive VR training, n=16	Control group Conventional VR training, n=15
Sex		
Men	4 (25%)	4 (27%)
Women	12 (75%)	11 (73%)
Age in years, mean (range)	28.6 (26-35)	29.6 (27-38)
Handedness		
Right	14 (88%)	15 (100%)
Left	1 (6%)	0 (0%)
Ambidextrous	1 (6%)	0 (0%)

Table 2. Description of the 360° videos played as backdrop during the procedures in the immersive VR environment

	Theme	Description
Video 1 (57 s)	Calm	The staff stays at their positions and organize their area of responsibility in a quiet fashion; e.g., the scrub nurse rearranges instruments on the table
Video 2 (59 s)	Calm	Same as video 1, except the floor nurse walks behind the OR tower to a closet and back
Video 3 (48 s)	Light stressor	The anesthesiologist enters and conducts a conversation with the nurse anesthetist while the floor nurse simultaneously answers a ringing phone
Video 4 (59 s)	Severe stressor	A bleeding (2 ml/s) in the simulated procedure (different points of bleeding depending on procedure number) is triggered at the start of the video (t = 221 s). The surgical nurse comments on the patients' bleeding and tells the floor nurse to call the senior surgeon on duty; the surgical assistant orders for equipment for conversion to open surgery to be available; the nurse anesthetist orders a transfusion pack to be available; and the nurse anesthetist calls the anesthesiologist and relays the situation in the OR

Table 3. Effects on relative reaction time

	Estimated marginal means	95% CI	p
Immersive VR	1.66	1.63-1.69	<0.001
Conventional VR	1.58	1.55-1.61	
Repetition 1	1.69	1.65-1.72	<0.001
Repetition 2	1.61	1.57-1.65	
Repetition 3	1.56	1.52-1.60	
Not cauterizing	1.59	1.56-1.61	<0.01
Cauterizing	1.65	1.61-1.69	

Table showing higher CL in the IVR group compared to the CVR group and higher CL while cauterizing. Both groups lowered their CL at the same rate during the three procedures.

Table 4. Simulator performance metrics

	Immersive VR	Conventional VR	
	Estimated marginal means (95% CI)	Estimated marginal mean (95% CI)	p
Total time (s)	533 (492-575)	409 (368-450)	<0.001
Blood loss (ml)	190 (160-221)	140 (110-169)	=0.02
Diathermy damage	2.3 (1.8-2.8)	1.3 (0.8-1.8)	<0.01
Tube cut: distance from uterus (mm)	4.5 (3.2-5.7)	4.6 (3.4-5.8)	ns
Major vessel cut (#)	0.14 (0.03-0.26)	0.10 (-0.02-0.21)	ns
Path length right instrument (m)	9.7 (8.8-10.6)	6.4 (5.5-7.3)	<0.001
Path length left instrument (m)	5.2 (4.7-5.8)	3.0 (2.4-3.5)	<0.001
Angular path right instrument	1670 (1480-1860)	956 (769-1143)	<0.001
Angular path left instrument	943 (836-1050)	508 (403-613)	<0.001

ns = not significant