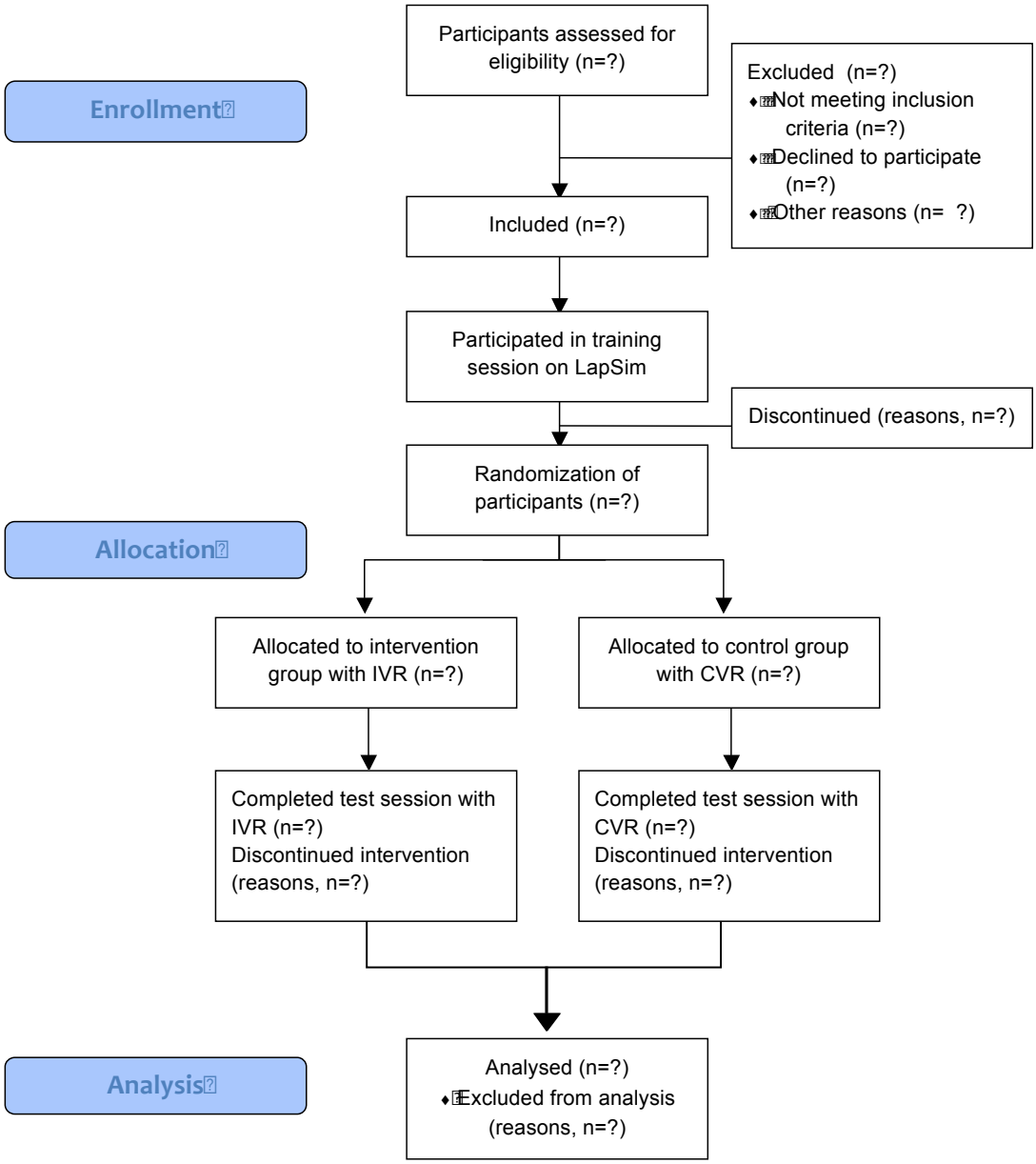
Title

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

Trial flow chart

This trial follows the CONSORT-statement for randomized trials.



List of abbreviations

CAMES: Copenhagen Academy for Medical Education and Simulation

CLT: Cognitive Load Theory

CVR: Conventional Virtual Reality

HMD: Head Mounted Display

IVR: Immersive Virtual Reality

VR: Virtual Reality

Background

During the last three decades, minimally invasive surgery and laparoscopic procedures have become widespread due to several advances compared to conventional, open surgery: such as a reduction of the surgical trauma, shorter hospitalization and faster postoperative recovery[1]. However, laparoscopic surgery requires other skills than open surgery, partly because the surgeon has to overcome the fulcrum effect, the limited degrees of freedom and the loss of depth perception due to 2D projection. These psychomotor skills need to be learned and trained, preferably, in a risk-free environment such as virtual reality (VR) simulation or other simulation-based training.

VR simulation is increasingly implemented in laparoscopic skills training programs, as it is well known to improve surgical trainees' operating performance and decrease their operating time[2].

Opponents of VR simulation often raise concerns that there is a substantial gap between the simulation-environment and the real-life operating room[3]. The effect of this being that skills obtained in VR simulation may not be transferable to the operating room. New technology may improve this by using VR head-mounted displays (HMDs) in combination with VR laparoscopic simulators. The result is an immersive VR (IVR) simulation experience where the surrounding environment from the real OR is integrated in the simulation, which may increase the transferal of skills as the environment in the simulation approaches the real-life environment.

An increase in the realism of simulations, e.g. IVR, might also increase the strain on the trainees' working memory and might lead to an overload of the trainees' cognitive abilities. This, according to cognitive load theory (CLT), will inhibit learning and a measurement of cognitive load is therefore of value when experimenting with IVR-training. CLT is used to optimize simulations-based training programs by examine how the cognitive load affects the actual learning. CLT is based on an information-processing model of the human cognitive architecture[4], which posits that information sensed from the environment must be processed by working memory before it can be consolidated and stored in long-term memory in the form of schemas[4]. Most literature on CLT considers three different types of cognitive load:

- Intrinsic load is a direct function of the complexity of the performed task and the expertise of the learner.
- Extraneous load is a result of superfluous processes that do not directly contribute to learning.
- Germane load is caused by learning processes that deal with the intrinsic load such as schema formation[4]

The different sources of cognitive loads are mainly considered additive, and if the sum surpasses working memory capacity, it can result in overloading. Such

cognitive overload hinders learning as excessive intrinsic and extraneous load takes up capacity from the germane load (i.e. actual learning). A goal of educational designers is to optimize germane load and decrease extraneous load.

Up until now, investigations of the use of IVR simulation in medical education are meagre. Huber et al.[5] found IVR simulation of laparoscopy to be clinical and technical feasible. Interestingly, they report that participants had longer operating time and higher error rates during IVR than during VR simulation of the same tasks. In the framework of CLT, a plausible explanation for this finding could be an increased extraneous load of the IVR simulation environment compared with conventional VR simulation. In another study, a significantly higher cognitive load was found in the more complex dissection environment compared with VR simulation for the mastoidectomy procedure[6]. To our knowledge, no study has previously investigated the role of cognitive load in IVR laparoscopic training.

We hypothesize that IVR increases cognitive load compared with conventional virtual reality (CVR) simulation of laparoscopic surgery. In this study, we examine the cognitive load and performance of a laparoscopic procedure in IVR and CVR in a randomized, controlled setup. The results will potentially have implications for organization of future training.

Primary objective
To compare immersive virtual reality (IVR) simulation training using a head-mounted display with virtual reality laparoscopic simulation training on LapSim (CVR) with regards to cognitive load.

Secondary objectives
To compare immersive virtual reality (IVR) simulation training using a head-mounted display with virtual reality laparoscopic simulation training on LapSim (CVR) with regards to performance

To investigate side effects of IVR simulation training such as motion sickness, feeling of immersion etc.

Setting and time
The study will be conducted at the Simulation Centre at Copenhagen Academy for Medical Education and Simulation, Rigshospitalet, Denmark, from February to May 2018.

Participants and equipment
Participants are recruited by inviting newly graduated doctors.

Inclusion criteria

- Obtained medical license
- Residents working in Denmark
- Signed informed consent

Exclusion criteria

- Previous participation in trials involving laparoscopic training
- Having performed one or more laparoscopic procedures as primary surgeon, including supervised procedures
- No informed consent
- Does not speak Danish on a conversational level

Participant discontinuation and withdrawl

A participant, who no longer wishes to participate in the trial, can withdraw his/her informed consent at any time. Analyses will be performed according to the intention-to-treat principle where all randomised participants will be analysed. Therefore it is in the interest of the trial to collect as much data from each participant as possible, so the amount of 'missing data' is minimized. If a participant withdraws, the investigator will ask if his/her data may be used in the analyses. Only if the participant specifies, his/her data cannot be used in the analyses, will a new person be randomised. The investigator can discontinue a participant from the trial if the participant does not comply with the training protocol.

Study design

All participants will first complete a questionnaire on background demographics.

Next, all participants will receive a 1-hour of introduction to the laparoscopic simulator, where they will train four different basic skills tasks (see appendix for descriptions). They will train 12 minutes on each basic skill. Subsequently, they will be introduced to the salpingectomy with an ectopic pregnancy procedure and complete one supervised attempt of this procedure on the LapSim (CVR).

Then the participant will be randomized to continue training in either the HMD-VR (IVR) or the LapSim (CVR) simulation environment. After randomization, participants will perform three attempts of the salpingectomy with an ectopic pregnancy procedure without any guidance. The participant will be allowed a maximum of 20 minutes pr. Attempt.

During the simulations, the cognitive load will be estimated using an external secondary-task measurement of reaction time in hundredths/s using a commercially available reaction timer (American Educational Products LLC, USA) where participants responds to an auditory cue by pressing a foot pedal. Reaction time will be measured in series of four repeated measurements: 1) before and

after the procedure (baseline) and 2) at four different points in time during the simulation, representing two calm phases and two phases with stressors during the IVR simulation and at three calm phases and one phase with a stressor during the CVR simulation.

The IVR simulation will consist of four different videos. The actors in these videos starts and ends in the same positions allowing the videos to be played in any order after one another.

Setup: The assistant to the left is holding the camera, the surgical nurse is opposite and a bit to the left, the floor nurse is at the end of the operating table which is further to the left and the nurse anesthetist is to the right behind the anesthetic cover. The anesthesiologist is not in the operating theatre.

Video 1 (57 seconds) and video 2 (59 seconds):

There will be no talking and the surgical staff will just move slightly around and/or watch the surgeon.

Video 3 (48 seconds):

The anesthesiologist enters and conducts a conversation with the nurse anesthetist then leaves. Simultaneously the phone rings and the floor nurse will be talking on it.

Video 4 (59 seconds):

The surgical nurse comments that the patients' bleeding has increased., then everyone in the operating theatre will be giving orders to the floor nurse while she is on the phone. Simultaneously the nurse anesthetist will be calling the anesthesiologist.

The sequence of the videos will be:

Sequence of videos	1	2	3	1	4	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2

Which will amount to a total of 1.208 seconds or 20 minutes and 8 seconds, but we expect most of our participants to be done with the procedure at around 8 minutes.
The cognitive load will be estimated at t=80 (calm), t=130 (stressor), t=180 (calm) and t=240 seconds (stressor). At t=228 seconds a bleeding close to the ovarium will be triggered using the TeamSim software. This corresponds with the comment from the surgical nurse that the patient is bleeding more.
Cognitive load testing times will be the same for the CVR simulation and the bleeding will be triggered at the same time as well. The bleeding will act as the stressor at the cognitive load estimate at t=240 seconds for the CVR.

Illustration of video sequence and cognitive load measurements for IVR for the first 339 seconds:

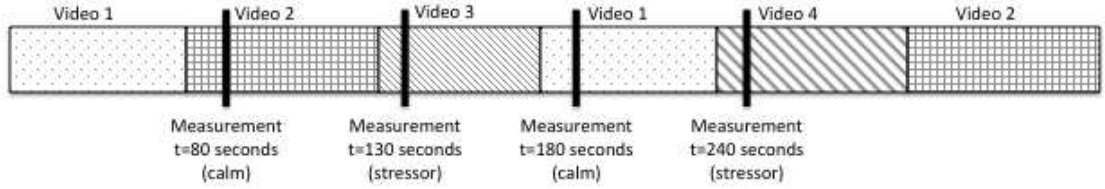
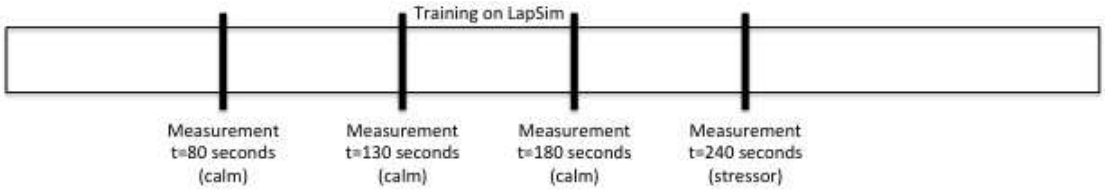


Illustration of cognitive load measurements for CVR for the first 339 seconds:



Randomization and blinding

The participants will be randomized to either group A (IVR) or group B (CVR), using the service Sealed Envelope[tm]. The allocation sequence is computer-generated with a varying block size kept concealed from the investigators. Allocation of participants will be performed after one supervised salpingectomy with an ectopic pregnancy procedure on LapSim.

One stratification variable is used for randomisation: Sex (man/woman). The stratification variable is included because it may have an impact on initial laparoscopic simulator performance according to previous published studies.[7-9]

Participants will be blinded with regards to simulator metrics. Blinding of participants or data collector to the allocation/intervention is not possible. During the statistical analysis, data regarding group and participant will be pseudonymized, so that the person responsible for analysis (SA) will not be aware which group is the control group and which group is the intervention group.

Outcome measures

Secondary reaction time will be used as an estimate of cognitive load and is measured in hundredths second, with a maximal reaction time of 99/100 second. Reaction time during simulation will be calculated relatively to the mean initial baseline reaction time of the individual participant.

The simulator will automatically gather data regarding the participants' performance, measuring:

Total Time (s) Blood Loss (ml)

Pool Volumen (ml)	Ovary Diathermy Damage (s)
Tube Cut: Uterus Distance (mm)	Bleeding Vessel Cut (max 1)
Left Instrument Path Length (m)	Right Instrument Path Length (m)
Left Instrument Angular Path (degrees)	Right Instrument Angular Path (degrees)

The background questionnaire will include questions on demographics, prior simulation training, prior laparoscopy experience, experience with HMD and video game experience. Finally, a questionnaire regarding motion sickness[10] and the feeling of being immersed during the training will be provided after IVR and CVR simulation training.

Data collection

The primary investigator will collect data, and the data will be pseudo-anonymized and kept in such a way that collected data will be stored separately from the anonymization key.

Statistics/Data analysis

Data will be analyzed using statistical software (SPSS) and relevant statistical methods (linear mixed models). Groups will be compared using independent samples t-test. The level of statistical significance will be set at alpha= 0.05.

Sample size

In general sample size calculations are based on the primary outcomes of 1) relative reaction time and 2) performance metrics, but due to repeated measurements there is no standardized way to estimate the number of participants. Based on similar studies, the authors suggest a total of 48 participants in the trial will be sufficient to detect a difference in reaction time of 5 % and performance of 10 %.

Economic

No external funding has been received.

Ethical considerations

Ethical approval or exemption will be obtained from the Regional Ethical Committee of the Capital Region, Denmark.

Written consent for the participation in the trial will be obtained.

The data involves no patient related information and approval by the Danish Data Protection Agency is not relevant.

The trial will be reported to a clinical trial database.

Publication
The findings of this trial are to be published in a peer-reviewed journal within the field of surgery and/or medical education.

Regardless of the outcome of the trial, the results will be published. If this is not possible through a scientific publication, a report will be compiled and made available online.

The sequence of authors on the publication will be as follows: Joakim Grant Frederiksen, Stine Maya Dreier Sørensen, Lars Konge, Morten Bo Søndergaard Svendsen, Morten Nobel-Jørgensen, Flemming Bjerrum and Steven A. W. Andersen.

Study group
Joakim Grant Frederiksen, BSc of Medicine, University of Copenhagen, Denmark, and Copenhagen Academy for Medical Education and Simulation, Center for HR, Region H, Denmark.

Stine Maya Dreier Sørensen, BSc of Medicine, University of Copenhagen, Denmark, and Copenhagen Academy for Medical Education and Simulation, Center for HR, Region H, Denmark.

Lars Konge, MD, PhD, Professor, Copenhagen Academy for Medical Education and Simulation, Center for HR, The Capital Region of Denmark

Morten Bo Søndergaard Svendsen, M.Sc.Eng., PhD, Copenhagen Academy for Medical Education and Simulation, Center for HR, Region H, Denmark

Morten Nobel-Jørgensen, M.Sc.Games., PhD, Center for Computer Games Research, IT University of Copenhagen, Denmark.

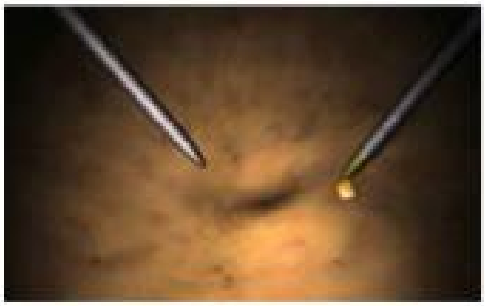
Flemming Bjerrum, MD, PhD, Dept. of Surgery, Herlev Gentofte Hospital, Capital Region of Denmark, Copenhagen, Denmark

Steven A. W. Andersen, MD, PhD. The Simulation Centre, Copenhagen Academy for Medical Education and Simulation, Center for HR, Region H, Denmark, and Dept. of Otorhinolaryngology-Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark.

References

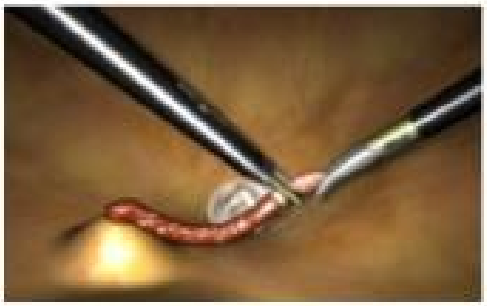
1. Sorensen, S.M., Savran, M.M., Konge, L. & Bjerrum, F. Three-dimensional versus two-dimensional vision in laparoscopy: a systematic review. *Surg Endosc* **30**, 11-23 (2016).
2. Nagendran, M., Gurusamy, K.S., Aggarwal, R., Loizidou, M. & Davidson, B.R. Virtual reality training for surgical trainees in laparoscopic surgery. *The Cochrane database of systematic reviews*, CD006575 (2013).
3. Moorthy, K., Munz, Y., Adams, S., Pandey, V. & Darzi, A. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. *Ann Surg* **242**, 631-639 (2005).
4. van Merrienboer, J.J. & Sweller, J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ* **44**, 85-93 (2010).
5. Huber, T.*, et al.* New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff. *Surg Endosc* **31**, 4472-4477 (2017).
6. Andersen, S.A., Mikkelsen, P.T., Konge, L., Caye-Thomasen, P. & Sorensen, M.S. Cognitive Load in Mastoidectomy Skills Training: Virtual Reality Simulation and Traditional Dissection Compared. *J Surg Educ* **73**, 45-50 (2016).
7. Strandbygaard, J.*, et al.* Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Ann Surg* **257**, 839-844 (2013).
8. Thorson, C.M., Kelly, J.P., Forse, R.A. & Turaga, K.K. Can we continue to ignore gender differences in performance on simulation trainers? *Journal of laparoendoscopic & advanced surgical techniques. Part A* **21**, 329-333 (2011).
9. White, M.T. & Welch, K. Does gender predict performance of novices undergoing Fundamentals of Laparoscopic Surgery (FLS) training? *Am J Surg* **203**, 397-400; discussion 400 (2012).
10. Gianaros, P.J., Muth, E.R., Mordkoff, J.T., Levine, M.E. & Stern, R.M. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine* **72**, 115-119 (2001).

Basic skills tasks on LapSim



Basic skill 1: Touch the gallstone with the tip of the right instrument before it disappears. Continue touching the gallstones as they appear while alternating between the right and left instrument.

Basic skill 2: Stretch the object until it releases from the tissue, move it to the target area, and release it when the target area changes color. Continue to move the object with the right and left grasper alternately until the exercise is completed.



Basic skill 3: Grasp the vessel in the highlighted area and stretch it until a new segment becomes highlighted. Move the ultrasound scissors to this highlighted area and press the pedal to cut the vessel. Move the excised segment to the target area and release when it changes color.

Basic skill 4: Start by grasping the top of the vessel package. Stretch the vessel and by carefully using the hook, catch the small thin vessels and stretch them before burning, so no burning damage will occur on the big vessels.

Questionnaire regarding motion sickness:

MOTION SICKNESS ASSESSMENT QUESTIONNAIRE (MSAQ)

Instructions: Using the scale below, please rate how accurately the following statements describe your experience

Not at all Severely

1----2----3----4----5----6----7----8----9

1. I felt sick to my stomach
2. I felt faint-like
3. I felt annoyed/irritated
4. I felt sweaty
5. I felt queasy
6. I felt lightheaded
7. I felt drowsy
8. I felt clammy/cold sweat
9. I felt disoriented
10. I felt tired/fatigued
11. I felt nauseated
12. I felt hot/warm
13. I felt dizzy
14. I felt like I was spinning
15. I felt as if I may vomit
16. I felt uneasy

MOTION SICKNESS ASSESSMENT QUESTIONNAIRE (MSAQ).

Instructions. **Using the scale below, please rate how accurately the following statements describe your experience**

Not at all Severely

1——2——3——4——5——6——7——8——9

1. I felt sick to my stomach (G)	9. I felt disoriented (Q
2. I felt faint-like (C)	10. I felt tired/fatigued (S)
3. I felt annoyed/irritated (S)	11. I felt nauseated (G)
4. I felt sweaty (P)	12. I felt hot/warm (P)
5. I felt queasy (G)	13. I felt dizzy (C)
6. I felt lightheaded (C)	14. I felt like I was spinning (C)
7. I felt drowsy (S)	15. I felt as if I may vomit (G)
8. I felt clammy/cold sweat (P)	16. I felt uneasy (S)

Note. G; Gastrointestinal; C: Central; P: Peripheral; SR; Sopite-related.

The overall motion sickness score is obtained by calculating the percentage of total points scored: (sum of points from all items/144) × 100. Subscale scores are obtained by calculating the percent of points scored within each factor: (sum of gastrointestinal items/36) × 100; (sum of central items/45) × 100; (sum of peripheral items/27) × 100; (sum of sopite-related items/36) × 100.