Cognitive load in mastoidectomy skills training: virtual reality simulation and traditional dissection compared

Steven Arild Wuyts Andersen, MD, Peter Trier Mikkelsen, MSc,† Lars Konge, MD, PhD,‡ Per Cayé-Thomasen, MD, DMSc* and Mads Sølvsten Sørensen, MD, DMSc**

*Department of Otorhinolaryngology—Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark; †Computer Graphics Lab, the Alexandra Institute, Aarhus, Denmark; and ‡Center for Clinical Education, Center for HR, The Capital Region of Denmark, Copenhagen, Denmark.

Correspondence: Steven A. W. Andersen, MD, Department of Otorhinolaryngology—Head & Neck Surgery, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen Ø, Denmark. Phone: 0045 20612006. E-mail: stevenarild@gmail.com.

Conflicts of interests: None.

Financial disclosures:
Steven A. W. Andersen has received an unrestricted grant from the Oticon Foundation for PhD studies. The development of the Visible Ear Simulator software has been financially supported by the Oticon Foundation.

Full citation: Andersen SA, Mikkelsen PT, Konge L, Cayé-Thomasen P, Sørensen MS. Cognitive load in mastoidectomy skills training: virtual reality simulation and traditional dissection compared. J Surg Educ. 2016 Jan-Feb; 73(1):45–50.
DOI: 10.1016/j.jsurg.2015.09.010

OBJECTIVE: The cognitive load (CL) theoretical framework suggests that working memory is limited, which has implications for learning and skills acquisition. Complex learning situations such as surgical skills training can potentially induce a cognitive overload, inhibiting learning. This study aims to compare CL in traditional cadaveric dissection training and virtual reality (VR) simulation training of mastoidectomy.

DESIGN: A prospective, crossover study. Participants performed cadaveric dissection before VR simulation of the procedure or vice versa. CL was estimated by secondary-task reaction time testing at baseline and during the procedure in both training modalities.

SETTING: The national Danish temporal bone course.

PARTICIPANTS: 40 novice otorhinolaryngology residents.

RESULTS: Reaction time was increased by 20 % in VR simulation training and 55 % in cadaveric dissection training of mastoidectomy compared to baseline measurements. Traditional dissection training increased CL significantly more than VR simulation training ($p<0.001$).

CONCLUSIONS: VR simulation training imposed a lower CL than traditional cadaveric dissection training of mastoidectomy. Learning complex surgical skills can be a challenge for the novice and mastoidectomy skills training could potentially be optimized by employing VR simulation training first because of the lower CL. Traditional dissection training could then be used to supplement skills training after basic competencies have been acquired in the VR simulation.

KEY WORDS: cognitive load, temporal bone dissection, mastoidectomy, virtual reality simulation, surgical skills training

INTRODUCTION

Complex skills are needed in temporal bone surgery and need to be taught to residents: 1) precise motor skills in handling drill, suction/irrigation and the operating microscope and 2) a deep knowledge and understanding of the anatomy and surgical relations of the temporal bone. This has traditionally been taught through cadaveric temporal bone dissection followed by supervised surgery. For the novice, novel and unorganized information and technical skills represent a complex learning task and surgical skills training should be organized to provide efficient learning.

During the last decade, simulation training has been increasingly employed in health professional education and also in surgical education supported by the development of virtual reality (VR) surgical simulators. Maintaining facilities for traditional cadaveric dissection training is very costly especially if trainees are provided access to "open lab" facilities, where surgical procedures can be trained at the convenience and needs of the individual trainees. However optimal from a surgical educational point, such facilities require constant care and supervision including sanitization of equipment, disposal of biological waste, acquisition and maintenance of necessary equipment. In addition, cadaveric temporal bones are becoming increasingly difficult to obtain due to regulation and safety issues. VR simulation training of mastoidectomy[1–4] can provide the opportunity for novices to acquire some of the necessary basic competencies in a safe environment before proceeding to other training modalities such as cadaveric dissection.

In complex learning tasks such as learning the mastoidectomy procedure, the learners' cognitive load (CL) in the learning situation should be considered because this could have implications for the organization of skills training. It could be hypothesized that because VR simulation is less complex it reduces the CL of the learner, in turn leading to better learning and making VR simulation training well suited for the initial training of novices. However, the haptic interaction and different visual cues in VR simulation could impose additional CL compared with other training modalities. This knowledge could instigate relevant changes in both the specific training modalities (i.e. virtual and cadaveric dissection training) and in the organization of temporal bone surgical skills training in general.

Cognitive load theory (CL-theory) is one of the leading theories of learning and provides a theoretical framework of the cognitive architecture with the basic assumption that working memory and information processing is limited.[5] The theory suggests three sources of cognitive load (CL) in any learning situation: 1) the intrinsic load of the learning task, 2) the extraneous load provided by the learning situation and 3) the germane load of the learning process itself.[6] Actual learning and skills acquisition can according to the theory be impeded if the total CL in the learning situation results in a cognitive overload, i.e. the limits in working memory and information processing are exceeded. Learning tasks can impose an extraneous CL resulting in a cognitive overload for example when the learner needs to integrate complex multisource information and psychomotor skills. This can be managed by employing instructional strategies and design principles to lower the extraneous load and optimize intrinsic and germane loads[6], leading to more efficient learning.

CL can be measured by a variety of methods including self-reported invested mental effort or difficulty, or by objective measures such as functional magnetic resonance imaging (fMRI) or the dual-task paradigm.[7] In the dual-task paradigm, performance in a secondary task estimates CL for example by measuring reaction time in response to a visual, auditory or tactile stimulus. Secondary task reaction time performance has been demonstrated to detect changes in CL in surgical skills training of novices.[8]

In this study, we wanted to compare CL in traditional dissection and VR simulation training of mastoidectomy using the dual-task paradigm with reaction time measurement for CL estimation.

MATERIAL AND METHODS

A total of 40 otorhinolaryngology residents (post-graduate year 2-5) participating in the national Danish temporal bone courses in January 2014 and 2015 were recruited for a study on cadaveric dissection and VR simulation training of mastoidectomy in a crossover study-design with participants receiving cadaveric dissection first in the 2015 course and VR simulation first in the 2014 course. During this study, we also performed the reaction time measurements during VR simulation and cadaveric dissection training presented here. Participants signed informed consent for participation.

Participants were novices regarding the mastoidectomy procedure with only limited mastoidectomy experience because participation in a temporal bone course is a prerequisite before beginning supervised temporal bone surgery.

Cadaveric dissection was set up with a cadaver head in a dissection tray, an operating microscope, an otosurgical drill and irrigation device operated by a pedal and a continuous vacuum suction (Figure 1A). Participants were given 60 minutes to complete a mastoidectomy up until the point of posterior tympanotomy without feedback or instructions by staff but guided by printed instructions. Participants were given longer time in cadaveric dissection than in VR simulation to compensate for the time needed to change drill bits and navigate the microscope and use the suction device, which was not needed in the VR simulator. Reaction time was measured manually (in hundredths/s) using a commercially available reaction timer (American Educational Products LLC, USA) where participants responded to an auditory cue (a beep) by pressing a pedal. Reaction time was measured before the procedure for baseline and three times during the procedure (t=5, t=25, and t=45 minutes).

In VR simulation training, participants were introduced to the simulator with a five-minutes hands-on exercise followed by three complete mastoidectomies up until the point of posterior tympanotomy. Participants were allowed 60 minutes for the first procedure and 30 minutes for the second and third procedure. We used the Visible Ear Simulator—a freeware VR temporal bone surgical simulator that can be downloaded from the Internet[9–10]—for VR simulation training. The simulator supports the Geomagic Touch™ (3D Systems, USA) haptic device for drilling with force-feedback (Figure 1B). A research version of version 1.3 of the simulator was developed to integrate a reaction time test for CL estimation. The reaction time test consisted of a secondary visual monitoring task where the participants had to press the key corresponding to a random letter displayed in a color-changing box (Figure 2). Reaction time was measured (in ms) before and after each VR procedure (baseline) and three times during the procedure (at t=5, 15, and 25 minutes) by the simulator software. Reaction times were normally distributed after log-transformation and Winsorized using two times the standard deviation as cut-off to counter extreme outlying values.

Data were analyzed with IBM SPSS Statistics version 22 for MacOS X (IBM Corp., Armonk, NY, USA). Simulation mean reaction times were calculated relative to the corresponding, individual session mean baseline measurement in order to compare the two different training and measurement modalities. Means and within-subjects effects were calculated using a linear mixed model for repeated measurements. Paired samples t-test was used to compare the means in VR simulation and cadaveric dissection. P-values below 0.05 were considered statistically significant.

The regional ethics committee (The Capital Region of Denmark) found this study to be exempt (H-4-2014-FSP 2).

RESULTS

Participants receiving cadaveric dissection training first and participants receiving VR simulation training first had comparable age (36 vs. 37 years, $p=0.26$), sex (55 % females vs. 50 % females, $p=0.76$), and years of training in otorhinolaryngology (3.8 vs. 4.4 years, $p=0.28$). No correlations where found between these factors and relative reaction time during cadaveric dissection or VR simulation training.

The reaction time was increased during the procedure in VR simulation and cadaveric dissection training compared to the individual baseline in each modality. The mean relative reaction time (unitless) was 1.55 (95 % confidence interval 1.52–1.59) in cadaveric dissection training and 1.20 (95 % confidence interval 1.18–1.22) in VR simulation training, meaning that reaction time increased by 55 % in cadaveric dissection training and 20 % in VR simulation training compared with baseline. This difference in relative reaction time between training modalities was found to be statistically significant ($p<0.001$).

Time of measurement during cadaveric dissection training (t=5, 25, and 45 minutes) was not found to be of statistical significance ($p=0.26$), meaning that the relative reaction time was stable during the cadaveric dissection training (Table 1). In contrast to this, time of measurement (t=5, 15, and 25 minutes) during VR simulation training was found to be statistically significant ($p<0.002$) with an increased reaction time in the later measurements (Table I).

Participants performed three procedures during VR simulation training. Procedure number was not found to impact on relative reaction time in VR simulation training ($p=0.55$).

DISCUSSION

In this prospective, crossover study on CL in traditional cadaveric dissection and VR simulation training, we used reaction time as an estimate for CL and found that the relative reaction time was significantly increased during both cadaveric dissection and

VR simulation training. Cadaveric dissection training increased CL significantly more than VR simulation training.

We used the dual-task performance paradigm to estimate the CL because CL is a construct "describing the internal processes of information processing that cannot be observed directly".[7] Different methods for assessing CL have been described and one way of classifying the methods is in the 'objectivity' and 'causal relation' dimensions with the dual-task paradigm being categorized an objective and direct method.[7] The dual-task paradigm can principally be employed in two ways: [7] the secondary task is added to the primary task in order to increase the load of the learning task—performance on the primary task being the variable of interest—or by using the secondary task to measure the load induced by the primary task—performance on the second task reflecting to the varying load of the primary task. We chose the latter method because our participants obviously needed to learn the mastoidectomy procedure and not be unnecessarily distracted by the secondary task. In addition, it has been demonstrated that dual-task reaction time with response to a vibrotactile stimulus was a sensitive measure of CL for novice learners in the initial stages of simulation-based surgical skills training.[8] In this study, participants were reaction time tested in response to a vibrotactile stimulus.

Few other studies have explored cognitive load in simulation-based surgical skills training even though surgical skills training represents a complex learning situation with a high risk of cognitive overload. One study on a VR simulator for salpingectomy/salpingotomy found that improvements in dexterity parameters (time and number of movements) following from repeated practice in the VR simulator were correlated with a decreasing CL measured using a subjective-indirect method (mental-effort questionnaires).[11]

This study is the first study comparing CL of the same surgical procedure in two different skills training modalities. We found that CL was significantly more increased during traditional cadaveric dissection training than VR simulation training. Within the framework of CL-theory, this could possibly be explained by a reduction of either the extraneous load of the learning situation or the intrinsic load of the learning task. Extraneous load can be caused by learning situations in which the learner has to apply weak problem-solving methods, lacks suitable guidance, has to integrate information sources or skills separated in time and place, or needs to search for necessary information.[6] Intrinsic load is dependent on the learning task and the number of elements to be processed simultaneously—the element interactivity—and can only be managed by modifying the task.[6] This could suggest that some of these mechanisms are causing VR simulation training to impose a lower CL than cadaveric dissection training of mastoidectomy for novice trainees.

The VR simulator presents the instructions to the procedure on-screen with a step-by-step instruction with text and simulator-images. This could contribute to a decreased extraneous load compared with cadaveric dissection training where information on the procedure was found in table desk manuals that cannot easily be consulted while wearing gloves and a surgical gown and where the appropriate information needs to be sought out. The VR simulator also reduces complexity of the procedure because navigation, drilling and change of drill burrs are performed using a single haptic device and the keyboard. In addition, the Visible Ear Simulator version 1.3 does not simulate bleeding or bone dust. This contributes to markedly lower element interactivity in VR simulation training than in traditional cadaveric dissection training where several instruments are handled simultaneously (drill, suction/irrigation, pedal, and operating microscope). Whereas it is difficult to reduce the element interactivity in traditional cadaveric dissection training other interventions such as modifying and simplifying tasks and instructions could potentially reduce CL and should be investigated.

We found the relative reaction time to be slightly increased at the later time points in VR simulation training. When the participants reach more complex stages near vital structures in the later phases of the procedure, the CL of the trainees may increase thereby explaining this finding. In contrast, we did not find the same pattern in cadaveric dissection training where CL was found to be equally increased compared to baseline at all time points. However, this could simply be the effect of the overall very high CL during cadaveric dissection training, making the minor changes in difficulty negligible.

One of the limitations of the present study is that CL was estimated using two different methods of measurement: in VR simulation training we used a visual cue integrated directly into the simulator for automated measurement and in cadaveric dissection training the cue was auditory and provided by a human-operated reaction timing device. The differences in scales were overcome by calculating the reaction time relative to individual baseline measurements in each modality. However, to estimate CL precisely, both the primary and secondary task needs to make demands on the same cognitive resources. The primary task (mastoidectomy) places heavier demands on visual working memory than on auditory working memory. This could lead to an underestimation of the actual CL during the cadaveric dissection training and consequently to an underestimation of the difference in CL between VR simulation and cadaveric dissection training. The manual reaction-timing device could have been used to measure reaction time during VR simulation training or to establish normative data, mitigating the limitations of using different methods for measuring reaction time but this was not feasible in the context of this study. Another potential weakness of this study was that participants were not directly randomized to the order of cadaveric dissection training and VR simulation training. This was due to the practical organization of the course, but an equal number of participants received cadaveric dissection training before VR simulation training and vice versa, balancing out any effect of order of training on the CL means.

In this study, we did not examine the performance on the primary task and the purpose of this study was not to make any conclusion on the effect of CL or training modality on the actual learning outcome. It also remains without doubt that the virtual training platform provides a different learning experience from cadaveric dissection training. However, from CLT derives an array of possible interventions to lower the CL by modifying the design of instructions and the learning situation.[6] These could be applied in both VR simulation and traditional cadaveric and should be explored in future studies.

CONCLUSION

It is highly relevant to consider CL in surgical training especially in the context of novice learning: it can be beneficial for efficient learning to start simple and move towards more complex training once basic competencies have been acquired. Our findings suggest that VR simulation training provided the training platform in which the CL was lower. This could further support that VR simulation training should be employed first in mastoidectomy skills training and then succeeded by cadaveric dissection training to improve the trainee’s skills.

ACKNOWLEDGEMENTS:

None.

REFERENCES

1. Morris D, Sewell C, Barbagli F, Salisbury K, Blevins NH, Girod S. Visuohaptic simulation of bone surgery for training and evaluation, IEEE Comput Graph Appl. 2006;26(6):48–57.

2. Zirkle M, Roberson DW, Leuwer R, Dubrowski, A. Using a virtual reality temporal bone simulator to assess otolaryngology trainees. Laryngoscope. 2007;117(2):258–63.

3. O'Leary, S. J.; Hutchins, M. A.; Stevenson, D. R et al. Validation of a networked virtual reality simulation of temporal bone surgery. Laryngoscope. 2008;118(6):1040–6.

4. Wiet GJ, Rastatter JC, Bapna S, Packer M, Stredney D, Welling DB. Training otologic surgical skills through simulation-moving toward validation: a pilot study and lessons learned. J Grad Med Educ. 2009;1(1):61–6.

5. Sweller J. Cognitive load during problem solving: Effects on learning. Cogn Sci. 1988;12(2):257–285.

6. van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ. 2010;44(1):85–93.

7. Brünken R, Plass JL, Leutner D. Direct measurement of cognitive load in multimedia learning. Educational Psychologist. 2003;38(1):53–61.

8. Rojas D, Haji F, Shewaga R, Kapralos B, Dubrowski A. The impact of secondary-task type on the sensitivity of reaction-time based measurement of cognitive load for novices learning surgical skills using simulation. Stud Health Technol Inform. 2014;196:353–9.

9. Sorensen MS, Mosegaard J, Trier P. The visible ear simulator: a public PC application for GPU-accelerated haptic 3D simulation of ear surgery based on the visible ear data. Otol Neurotol. 2009;30(4):484–487.

10. Trier P, Noe KO, Sorensen MS, Mosegaard J. The visible ear surgery simulator. Stud Health Technol Inform. 2009;132:523–525.

11. Bharathan R, Vali S, Setchell T, Miskry T, Darzi A, Aggarwal R. Psychomotor skills and cognitive load training on a virtual reality laparoscopic simulator for tubal surgery is effective. Eur J Obstet Gynecol Reprod Biol. 2013;169(2):347–52.

Table I. Mean relative reaction times at different time points in VR simulation training and dissection training.

	Mean relative reaction time (unitless)	95 % Confidence Interval
Simulation		
5 minutes	1.15	1.12–1.19
15 minutes	1.22	1.19–1.26
25 minutes	1.23	1.20–1.27
Dissection		
5 minutes	1.59	1.49–1.69
25 minutes	1.56	1.47–1.65
45 minutes	1.52	1.42–1.61

Figure 1. A) An example of a set-up for traditional cadaveric temporal bone dissection training and B) the set-up for VR temporal bone simulation on a laptop with the Visible Ear Simulator and the Geomagic Touch™ haptic device.

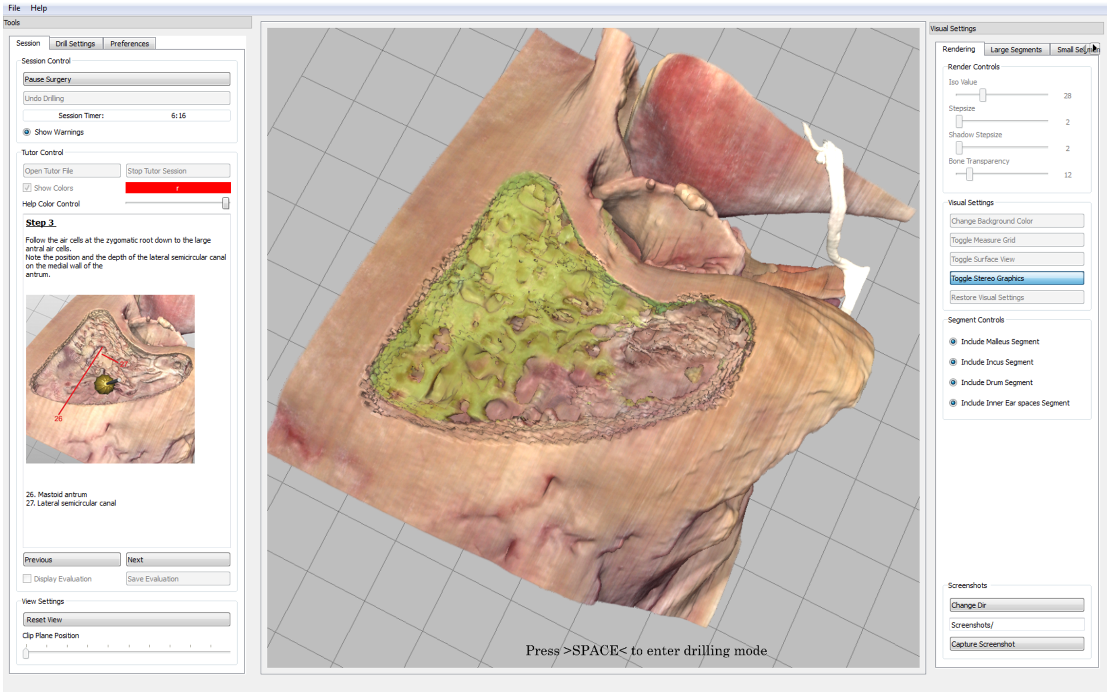

Figure 2. A screenshot from the experimental version of the Visible Ear Simulator. The reaction time test (red box) is found above the step-by-step instructions. The simulator-integrated tutor-function is greenlighting the volume corresponding to the current step of the procedure.