Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation

Steven Arild Wuyts Andersen[1], MD; Peter Trier Mikkelsen[2], MSc; Lars Konge[3], MD, PhD; Per Cayé-Thomasen[1], MD, DMSc; Mads Sølvsten Sørensen[1], MD, DMSc

1. Department of Otorhinolaryngology—Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark; 2. The Alexandra Institute, Aarhus, Denmark; 3. Center for Clinical Education, Center for HR, The Capital Region of Denmark, Copenhagen, Denmark.

Conflict of interests:

None.

Financial disclosures:

Dr. 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.

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

Full citation: Andersen SA, Konge L, Cayé-Thomasen P, Sørensen MS. Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation. Laryngoscope. 2016 Feb; 126(2):E74–9.

DOI: 10.1002/lary.25449

ABSTRACT

Objective: Cognitive load theory states that working memory is limited. This has implications for learning and suggests that reducing cognitive load (CL) could promote learning and skills acquisition. This study aims to explore the effect of repeated practice and simulator-integrated tutoring on CL in virtual reality mastoidectomy simulation.

Study Design: Prospective trial.

Methods: 40 novice medical students performed 12 repeated virtual mastoidectomy procedures in the Visible Ear Simulator: 21 completed distributed practice with practice blocks spaced in time and 19 participants completed massed practice (all practice in one day). Participants were randomized for tutoring with the simulator-integrated tutor function. CL was estimated by measuring reaction time in a secondary task. Data were analyzed using linear mixed models for repeated measurements.

Results: The mean reaction time increased by 37 % during the procedure compared with baseline demonstrating that the procedure placed substantial cognitive demands. Repeated practice significantly lowered the CL in the distributed practice group but not in massed practice group. In addition, CL was found to be further increased by 10.3 % in the later and more complex stages of the procedure. The simulator-integrated tutor function did not impact on the CL.

Conclusion: Distributed practice decreased CL in repeated virtual reality mastoidectomy training more consistently than was seen in massed practice. This suggests a possible effect of skills and memory consolidation occurring over time. To optimize technical skills learning, training should be organized as time-distributed practice rather than as a massed block of practice, which is common in skills-training courses.

Keywords: Cognitive load, distributed practice, massed practice, virtual reality simulation, technical skills training, mastoidectomy training.

Level of evidence: n/a

Introduction

Simulation is an efficient learning tool that can be used to train complex technical skills such as surgical procedures[1] and virtual reality (VR) has been established as an evidence-based tool for surgical technical skills training including in temporal bone surgery.[2–7] However, complex surgical tasks such as the mastoidectomy procedure—whether practiced in a VR environment or in a dissection lab—can provide a learning challenge for the novice.

Cognitive load theory (CL theory) as described by Sweller[8] in 1988 has become one of the dominant learning theories in medical education.[9] According to CL theory, working memory and the capacity for information processing is limited. Actual learning can be inhibited if the summed effect on cognitive load (CL) of the task itself (intrinsic load), the learning situation (extraneous load), and the learning process (germane load), exceeds the cognitive capacity of the learner resulting in a cognitive overload.[10] The simultaneous mental integration of novel and unorganized information and complex psychomotor skills can impose an extraneous CL. For the novice trainee, learning a surgical procedure such as mastoidectomy can therefore result in a cognitive overload that is detrimental to learning.

Optimizing intrinsic and germane loads and lowering the extraneous load could according to CL theory lead to better learning and skills acquisition. Several instructional strategies and design principles to accomplish this have been proposed.[10] Valid measurements of CL are essential to study the effect on CL of such interventions. A range of methods for estimating CL has been established in the literature.[11] One of these methods is the dual-task paradigm in which CL is estimated by measuring performance on a secondary task. Secondary-task reaction time has been demonstrated to be sufficiently sensitive to detect changes in CL in initial surgical skills training of novices.[12]

It has been reported that acquisition of complex psychomotor surgical skills is greater if practice is distributed over several sessions rather than massed as a single block of training.[13] This could relate to memory consolidation and the spacing of practice has also been demonstrated to benefit learning motor skills.[14] In the CL theoretical framework, the effect of repeated practice relates to the construction of mental schemas, optimizing the germane cognitive resources for dealing with the intrinsic CL of the procedure.[10] Simulator-integrated tutoring could be used to reduce the extraneous load by employing some of design principles based on CL theory—for example the ‘split attention’ and ‘redundancy’ principles.[10]

In this study, we wanted to investigate the effect on CL of organizing repeated practice of mastoidectomy in a VR temporal bone simulator as distributed or massed practice, estimating CL by measuring secondary-task reaction time. Our hypothesis was that distributed practice would lower CL more than massed practice because memory consolidation and schema construction would be more effective. We also wanted to investigate whether ongoing assistance by a simulator-integrated tutor function reduced CL as suggested by the design principles based on CL theory and we therefore employed a 2x2 study design.

Material and methods

VR simulation platform

The Visible Ear Simulator (VES) is a freeware[15] VR temporal bone surgical simulator, which runs on a PC with a GeForce™GTX graphics card (Nvidia®, USA) and supports the Geomagic Touch™(3D Systems, USA) haptic device for force feedback and intuitive drilling.[16–17] The simulator features 3D-stereo graphics, an integrated tutor function and a step-by-step tutorial for the mastoidectomy procedure. A special research version of the simulator version 1.3 was developed which allowed for individual user login with different preset conditions including a secondary reaction time task.

Participants

43 medical students from the Faculty of Health and Medical Science, University of Copenhagen, Denmark, volunteered for participation in this study, which was conducted in October and November 2013 (24 participants, distributed practice) and May 2014 (19 participants, massed practice). Previous training in VR temporal bone simulation was an exclusion criterion. Medical students at our institution have no exposure to temporal bone surgery during their preclinical or clinical studies. The participants only participated in one of the training programs and participation was an extracurricular activity.

21 of 24 participants completed the distributed training program and all 19 completed the massed training program and were included for study (participant characteristics in Table I). Three participants in the distributed group did not schedule further practice sessions after the initial training. Participants in the distributed group were significantly older (25.1 vs. 23.6 years), more often male (61.9 % vs. 26.3 %) and had a higher gaming frequency (2.3 vs. 1.6 on a 5-item Likert-like scale) and these factors were included as random effects in our analysis.

Primary and secondary task

The participants in both training programs received a brief lecture on the surgical anatomy of the temporal bone. The primary task was to perform a complete mastoidectomy with entry into the antrum and posterior tympanotomy. For the first session (pre-practice) the participants were allowed one hour and were instructed to explore the temporal bone in depth and learn the procedure. For the following 11 sessions the participants were allowed 30 minutes after which the simulator auto-saved the final-product for later manual analysis and closed. Preliminary analysis of the pre-practice mastoidectomy performance was performed using final-product analysis with one blinded rater using a modified Welling Scale.[18] During the procedure all the participants had access to the on-screen step-by-step guide with instructions and illustrations of the procedure but were otherwise self-directed.

The participants were given an unrelated secondary visual monitoring task in the simulator: they had to respond to the change of color in a box above the instructions by pressing the key corresponding to the letter that appeared within the box (Fig. 1). The participants were instructed to react as fast as possible while still performing the mastoidectomy. For each mastoidectomy, the reaction time test appeared four times in the two baseline measurements (before and after the procedure) and five times at three different predefined times during the procedure (at 5, 15, and 25 minutes). Reaction times were registered in milliseconds and auto-saved by the simulator.

Study design

A 2x2 study design was used to compare CL in distributed and massed practice with and without simulator-integrated tutoring (see flowchart, figure 2). Participants in both practice programs were randomized for simulator-integrated tutoring by computer-based randomization. Participants who received tutoring had the tutor function on during the first five procedures. The simulator-integrated tutor-function only green-lighted the volume to be drilled in each step corresponding to the on-screen guide available to both groups and the tutored group received no additional guidance or tutoring.

In the massed practice program, the participants performed all 12 sessions in one day. In the distributed practice program, the participants performed two complete simulations spaced at least three days apart (on average 7.7 days) from the next two complete simulations, totaling also 12 sessions.

Statistics

We calculated the mean simulation reaction times relative to the corresponding mean of the baseline measurements to get a measure of the relative change (unitless) in CL during simulation. The relative reaction time compensates for the between-subject variation (i.e. some participants are generally slower than others) and between-session variation (e.g. participants got tired with massed practice, sessions were on different days in the distributed group).

To define a reaction time for measurements where the participants were either too slow to react or missed the secondary task completely, the reaction time data were Winsorized using two times the standard deviation as cut-offs. The relative reaction time data for the 30-minute sessions (session 2-12) were analyzed with IBM SPSS Statistics version 22 for MacOS X (IBM Corp., Armonk, NY, USA) using a linear mixed model for repeated measurements. P-values below 0.05 were considered statistically significant.

Ethical approval

The regional ethics committee found that this study was exempt (H-4-2013-FSP-088).

Results

To address concerns regarding the differences in baseline characteristics between the two groups, we analyzed preliminary data on primary task performance. We found the two groups to have comparable mastoidectomy performances in session 1 (the pre-practice round) with the distributed group scoring 10.4/26 and the massed group 9.7/26 ($p=0.55$). In the distributed practice program, simulation practice was placed on different days according to the participant's individual schedule. This did not impact on the relative reaction time because no correlation was found between the number of days since last simulation and the change in relative reaction time between the time-spaced sessions (Pearson's r 0.061, $p=0.54$).

We found the overall mean reaction time to be significantly increased by 37 % during simulation compared with baseline ($p<0.001$).

The relative reaction time was used as the dependent variable in a linear mixed model with session number and pre-defined time of measurement during the procedure (t=5, 15 and 25 min) as repeated measurements. Practice program (distributed and massed), tutoring (tutor function on or off), session number and time of measurement as fixed variables, and age, sex and gaming frequency as covariates for random effects.

The initial model was reduced iteratively to fewest possible dimensions while retaining the highest degree of explanatory power. None of the covariates contributed significantly to the model and was excluded. Simulator-integrated tutoring was not found have an impact in the best model and was excluded. The final model included: 1) practice group by session number, and 2) the time of measurement during the procedure.

As our main result, we found that the relative reaction time decreased significantly with repeated practice for the distributed practice group (Fig. 3, left). This was not found to be the case for the massed practice group (Fig. 3, right).

The time of measurement during the procedure was found to have an effect on its own regardless of practice group (main effect): the relative reaction time was significantly lower in the beginning of the procedure (at 5 minutes) than later (at both 15 and 25 minutes) ($p<0.001$). The reaction time was found to be increased by 10.3 % during the later and more difficult parts of the procedure compared with the reaction time at 5 minutes.

Discussion

In this study on the CL in repeated practice of mastoidectomy in a VR simulator, we found that the reaction time was increased on average by 37 % during simulation compared with baseline but decreased significantly with repeated and time-distributed practice. In addition, we found that the later and more complex stages of the procedure further increased reaction time. Simulator-integrated tutoring did not influence reaction times and had no interactions with the practice program. Secondary task performance including the measurement of reaction time is an established direct/objective method to estimate the CL[11] and in line with this assumption, we will discuss our findings on change in reaction time as a change in the CL.

Time spacing and automaticity

CL in repeated practice of complex psychomotor surgical skills has previously been reported to correlate with the learning curve in distributed practice in a VR laparoscopic simulator for salpingectomy/salpingotomy surgery using a subjective measurement of CL.[19] The time spacing of practice could be important and impact on the CL regardless of fatigue occurring with massed practice because acquiring complex psychomotor surgical skills is dependent on consolidation over time[14]. Nevertheless, many skills-training courses are organized as massed practice, which is inefficient for optimal skills acquisition. In our study, the interval between every other session was on average 7.7 days and this inter-training interval seemed to be sufficient for the consolidation to occur; however the upper and lower limits and the optimal inter-training interval in surgical technical skills training has not been established and are still debated.[20]

Repeated practice can be expected to improve primary task performance as the learner gains competency and experience, leading to increased spare attentional resources also for the performance on the secondary task. In the terms of CL theory: the intrinsic load of the learning task decreases with repetition leaving more cognitive capacity for the learning process and schema construction (optimized germane load). It has previously been demonstrated that novices gain proficiency in VR surgical simulation after relatively few practice sessions—although much more practice is needed before performance on the secondary task improves.[21] In contrast to this, experts demonstrate the immediate ability to spare attentional resources for performance on the secondary task[20]. This could reflect experts' automaticity of the primary procedure and would require a substantial level of practice for novices to achieve. Automaticity and CL are intimately related; true automaticity will be reflected in a minimal CL.

Complexity and tutoring

We found the CL to depend on the stage of the procedure. The mastoidectomy procedure gets increasingly more complex and demanding as the trainee reaches vital structures in the temporal bone including the dura, the facial nerve, the inner ear and the ossicles. This was reflected directly in the relative reaction time at different times during each session: participants had a 10.3 % increase in reaction time during the later stages of the procedure. Secondary task reaction time thereby seems to be a sensitive monitor, reflecting the CL as it fluctuates during training of the procedure.

We expected that the simulator-integrated tutor function with volumetric green lighting could reduce the CL by the 'split attention' and 'redundancy' principles suggested by the CL theory[10], providing the trainee with one-source information and real-time instructions on the procedure. However, we found no difference in the relative reaction time between the tutored and the non-tutored group. During the trial we observed that the tutored participants used a great deal of effort to completely remove the volume green-lighted for removal by the tutor function (see figure 1). By experience this task can be very difficult to accomplish without collisions with the vital structures and the task for the tutored group could thereby be considered more challenging. It could be hypothesized that this challenge could add to the CL of the tutored participants and even out any positive effect of the simulator-integrated tutor function, possibly explaining why the CL remained unchanged with tutoring.

Limitations

A limitation to our study was that participants were not randomized to practice group—only to tutoring. The two practice groups had comparable experience with VR simulation in other specialties, general computer usage as well as clinical experience but had a different distribution of gender, age and gaming frequency. However, this was not found to influence relative reaction time in the repeated measurements linear mixed model. None of the participants had previously been exposed to temporal bone surgery or VR temporal bone surgery, as it is not part of their curriculum at any level and this was reflected in the preliminary analysis of their primary task performance, where the two groups had equal performances in their pre-practice mastoidectomy. Individual motivation for learning and the context and relevance of the learning experience should be acknowledged as possible limitations.

Another limitation of this study is that we currently lack detailed knowledge on the primary task performance. Even though distributed practice with time spacing of training sessions provides a lower CL, there is still a need to study performance on the primary task in order to establish whether the distributed practice provides better learning than massed practice in VR mastoidectomy simulation. The same goes for the simulator-integrated tutor function's effect on primary task performance. The current simulator-integrated metrics are of limited use[18] and manual rating with multiple raters would be needed.

Perspectives

Novices can acquire basic surgical technical skills in a VR simulator where the CL is potentially lower than other training modalities and the risk of cognitive overload is reduced. By further reducing the CL with repeated VR simulation training, it can be speculated that CL would start at a lower point when continuing on to practice in more demanding learning environments such as dissection training or the operating room. The transferability of CL remains an important relationship to be explored in future studies because it could have implications for the integration of VR simulation-based training in surgery. The application of different learning-interventions based on CL theory could improve instructional design and optimize learning in VR simulation training of surgical technical skills.

Conclusion

Distributed practice of mastoidectomy in a VR simulator decreased the CL more consistently than was seen with massed practice of the procedure. This could have implications for the organization of mastoidectomy skills training courses. Skills and memory consolidation is essential in complex psychomotor learning and this could also explain why CL only decreased significantly when practice sessions were distributed in time. Simulator-integrated tutoring did not as expected from CL theory-based instructional principles counter this. Other CL lowering interventions could be implemented in VR surgical simulation and potentially improve novice performance.

Acknowledgements

The authors would like to acknowledge Sebastian Roed Rasmussen from Center for Clinical Education, Copenhagen, Denmark, for his assistance in data collection.

References

1. Cook DA, Hatala R, Brydges R et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011;306:978-88.

2. Zirkle M, Roberson DW, Leuwer R, Dubrowski A. Using a virtual reality temporal bone simulator to assess otolaryngology trainees. Laryngoscope 2007;117:258-263.

3. Khemani S, Arora A, Singh A, et al. Objective skills assessment and construct validation of a virtual reality temporal bone simulator. Otol Neurotol 2012;33:1225-1231.

4. Zhao YC, Kennedy G, Yukawa K, et al. Can virtual reality simulator be used as a training aid to improve cadaver temporal bone dissection? Results of a randomized blinded control trial. Laryngoscope 2011;121;831-837.

5. Zhao YC, Kennedy G, Yukawa K, et al. Improving temporal bone dissection using self-directed virtual reality simulation: results of a randomized blinded control trial. Otolaryngol Head Neck Surg 2011;144:357-364.

6. Nash R, Sykes R, Majithi A et al. Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator, J Laryngol Otol 2012; 126:663-669.

7. Wiet GJ, Stredney D, Kerwin T, et al. Virtual temporal bone dissection system: OSU virtual temporal bone system: development and testing. Laryngoscope 2012; 122 Suppl 1:1-12.

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

9. Young JQ, Van Merrienboer J, Durning S, Ten Cate O. Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. Med Teach 2014;36:371-84.

10. van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ 2010;44:85-93.

11. Brünken R, Plass JL, Leutner D. Direct measurement of cognitive load in multimedia learning. Educational Psychologist 2003;38:53-61.

12. 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.

13. Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. Surg Endosc 2002;16:957-61.

14. Shea CH, Lai Q, Black C, Park JH. Spacing practice sessions across days benefits the learning of motor skills. Human Movement Science 2000;19:737-760.

15. The Visible Ear Simulator. Available at http://ves.cg.alexandra.dk/. Accessed February 12, 2015.

16. 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;30:484-487.
17. Trier P, Noe KO, Sorensen MS, Mosegaard J. The visible ear surgery simulator. Stud Health Technol Inform 2009;132:523-525.
18. Andersen SAW, Cayé-Thomasen P, Sølvsten Sørensen M. Mastoidectomy performance assessment of virtual simulation training using final-product analysis. Laryngoscope 2014;125:431-5.
19. 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;169:347-52.
20. Stefanidis D, Walters KC, Mostafavi A, Heniford BT. What is the ideal interval between training sessions during proficiency-based laparoscopic simulator training? Am J Surg 2009;197:126-9.
21. Stefanidis D, Scerbo MW, Sechrist C et al. Do novices display automaticity during simulator training? Am J Surg 2008;195:210-3.

TABLE 1. Participant characteristics.

	Distributed	Massed	p*
n	21	19	
Age, mean (years)	25.1	23.6	0.03
Gender			
Male	61.9 %	26.3 %	0.02
Female	38.1 %	73.7 %	
Years of study, mean	4.4	3.8	0.15
In pre-clinical years of study	28.6 %	26.3 %	0.88
In clinical years of study	71.4 %	73.7 %	
Has any previous VR simulation experience	23.8 %	36.8 %	0.38
Gaming frequency, mean (1-5 Likert like scale)	2.3	1.6	0.02
Computer usage, mean (hours)	20.6	18.3	0.60

*One-way ANOVA

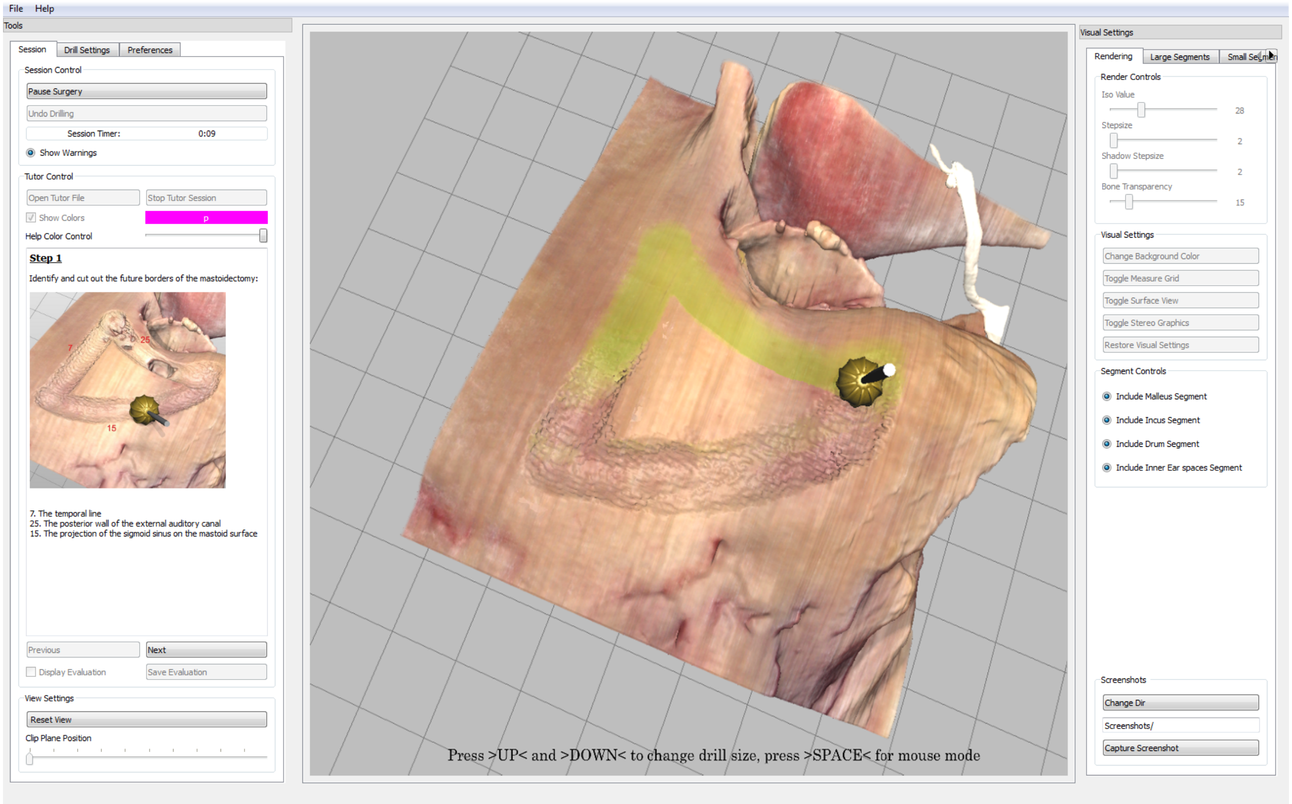

Fig. 1. Screenshot from the Visible Ear Simulator with the reaction time test (the pink box) to the left above the step-by-step tutorial. The first step of the mastoidectomy procedure is shown—in this case with the simulator-integrated tutor function on, green-lighting the volume to be drilled corresponding to the step in the on-screen tutorial on the left.

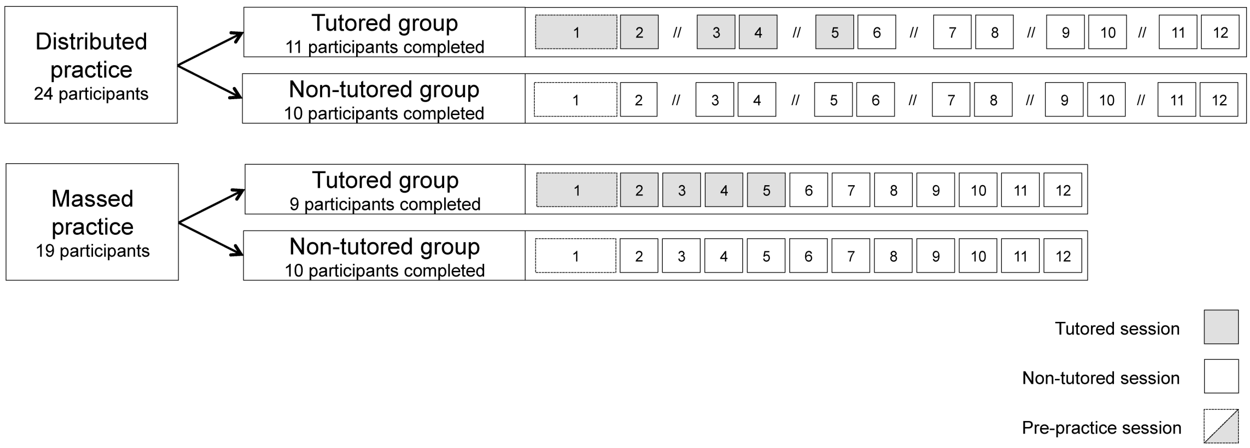

Fig. 2. Flow-chart. Participants in each practice group were randomized for additional simulator-integrated tutoring during the first five sessions. In distributed practice, every block of two repeated procedures were spaced by at least three days (on average 7.7 days). In massed practice, all of the repeated procedures were performed in the same day. Participants were allowed 30 minutes for the mastoidectomy procedure in each session following a 60-minute pre-practice session (session 1).

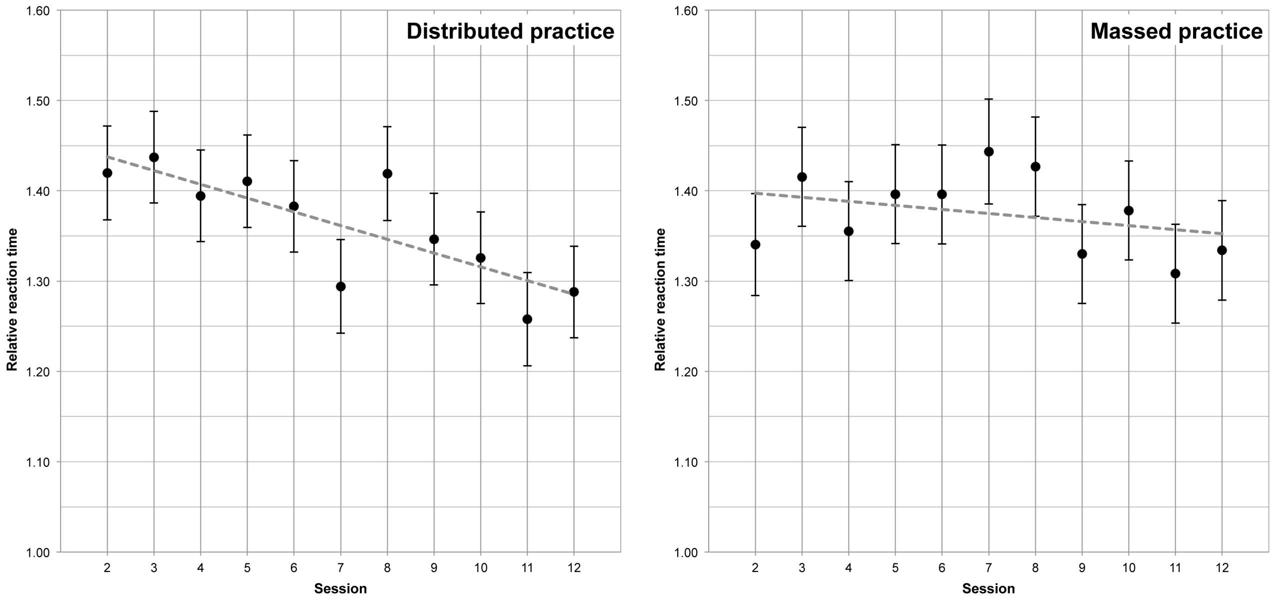

Fig. 3. Estimated marginal means of the linear mixed model for the relative reaction time by session in the distributed (left) and the massed practice group (right) with 95 % confidence intervals (vertical bars) and linear regression (dotted line).