Mapping the plateau of novices in virtual reality simulation training of mastoidectomy

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

1. Department of Otorhinolaryngology – Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark. 2. Copenhagen Academy for Medical Education and Simulation, Centre for HR, the Capital Region of Denmark, Copenhagen, Denmark. 3. Computer Graphics Lab, the Alexandra Institute, Aarhus, 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

Financial disclosures: Steven A. W. Andersen has received a grant from the Oticon Foundation for Ph.D. studies. The development of the Visible Ear Simulator software has been financially supported by the Oticon Foundation. The foundation played no role in the design or conduction of the study.

Full citation: Andersen SA, Konge L, Mikkelsen PT, Cayé-Thomasen P, Sørensen MS. Mapping the plateau of novices in virtual reality simulation training of mastoidectomy. *Laryngoscope*. 2017 Apr;127(4):907–914.

DOI: https://dx.doi.org/10.1002/lary.26000

Objectives: To explore why novices' performance plateau in directed, self-regulated VR simulation training and how performance can be improved.

Study-design: Prospective study.

Methods: Data on the performances of 40 novices who had completed repeated, directed, self-regulated VR simulation training of mastoidectomy were included. Data were analyzed to identify key areas of difficulty and the procedures terminated without using all the time allowed.

Results: Novices had difficulty in avoiding drilling holes in the outer anatomical boundaries of the mastoidectomy and frequently made injuries to vital structures such as the lateral semicircular canal, the ossicles and the facial nerve. The simulator-integrated tutor-function improved performance on many of these items but over-reliance on tutoring was observed. Novices also demonstrated poor self-assessment skills and often did not make use of the allowed time, lacking knowledge on when to stop or how to excel.

Conclusions: Directed, self-regulated VR simulation training of mastoidectomy needs a strong instructional design with specific process goals to support deliberate practice because cognitive effort is needed for novices to improve beyond an initial plateau.

Key-words: virtual reality simulation, temporal bone surgery, directed self-regulated learning, self-assessment, surgical training

Level of evidence: n/a.

INTRODUCTION

Surgical education has undergone a paradigm shift due to technological advances in surgery such as laparoscopy, restricted working hours of both trainees and supervisors, and an increased focus on patient safety. This has propelled simulation-based education and the development of advanced virtual reality (VR) simulators for technical skills and procedural training, allowing parts of surgical training to occur independently of patients, service duties, and supervision, in addition to supporting individual training needs.[1]

Cadaveric dissection has always been the best option for practicing temporal bone surgery in preparation for supervised surgery.[2] However, dissection training is costly and needs appropriate facilities and supportive staff in addition to human temporal bones, all of which are becoming increasingly scarce resources. Altogether, this limits training opportunities for the novice. VR simulation training is a new supplement to bridge the gap between these shortcomings and the need for skilled surgeons. In temporal bone surgery, multiple VR simulators have been developed and investigated.[3-8]

One of the potential benefits of simulation-based training of surgical skills is that it supports unsupervised or self-guided practice.[9] We have previously investigated the learning curves of directed, self-regulated mastoidectomy training in a VR temporal bone simulator:[10] although performance improved initially with repeated practice, the final-product score plateaued early and at a low score of 16 out of 26 points possible. Only a few participants reached a consistent performance level above this plateau suggesting a ceiling effect.[11] For mastoidectomy training in the OR, several milestones in the progression towards competency have been identified and the most difficult parts of the procedure required significant amounts of operative exposure, feedback and practice.[12] The complexity of the procedure could therefore also cause a limit on novice performance in directed, self-regulated training.

Although technical skills simulation can facilitate unsupervised learning, concerns are that trainees adopt bad habits and poor judgment.[13] Often, simulation-based training provide only limited guidance despite feedback being a critical component in surgical skills training.[14] Learning in the context of simulation-based training is therefore often dependent on accurate self-assessment. The ability to correctly self-assess varies the with the domain to be self-assessed but self-assessment of surgical technical skills seem to improve with repeated practice of the surgical task.[15] Literature also suggests there is a relation between self-assessed proficiency and the learning curve ceiling[16]. Other factors could also potentially modify novices' ability to self-assess thereby limiting their performance: The consolidation of skills occurring with repeated and distributed practice but not with massed practice,[10] or simulator-integrated tutoring, which might be supportive but also could lead to over-reliance.[10,17]

Evidence points to self-regulated learning in simulation-based training being effective.[18] However, especially for the novice, self-regulation can be difficult[19] and directed, self-regulated learning calls for a strong instructional design.[20] To improve directed, self-regulated VR simulation training of mastoidectomy, we therefore wanted to further analyze the performance of novices and to deconstruct the performance plateau and identify key areas of difficulty in VR simulation mastoidectomy. We also hypothesized that novices have limited self-assessment skills in relation to the mastoidectomy procedure and wanted explore this because of the profound implications for training.

MATERIAL AND METHODS

VR simulation platform

The Visible Ear Simulator is a VR temporal bone simulator based on high-resolution digital photos of cryo-sections of a fresh frozen human temporal bone.[7,21] The simulator is available as academic freeware[22] and runs on most newer personal computers with a high-end graphics card (i.e. Geforce GTX, Nvidia, Santa Clara, CA, USA). A Geomagic Touch (3D Systems, Rock Hill, SC, USA) haptic device is recommended for drilling with force feedback. Stereographics can be accomplished with anaglyphic 3D glasses. The simulator has a built-in step-by-step guide for mastoidectomy and also the option of intuitive visual guidance by the simulator-integrated tutor-function. The built-in guide is located in the left panel of the simulation software and resembles a traditional temporal bone manual and each step of the procedure is explained with a brief text, a picture from the simulator illustrating the step, and key anatomical landmarks indicated. The integrated tutor-function color-codes the volume of bone to be drilled in each step corresponding to the built-in guide and intuitively visualizes the volume to be removed directly on the interactive model in the central workspace.

Participants

For this study, we included data on the VR simulation performances of a total of 40 medical students from the University of Copenhagen, Denmark, who had previously completed mastoidectomy training on a VR temporal bone simulator with repeated practice in either a distributed condition (practiced blocks spaced by at least three days, 21 participants completed training) or in a massed condition (all practice completed in a single day, 19 participants completed training).[10,23] Participants were all novices regarding the procedure and training was organized as a voluntary extracurricular activity.

Study design

As previously detailed, the training program consisted of 12 repeated mastoidectomy procedures on an identical virtual temporal bone. Each practice cohort (distributed and massed) were randomized for support by the simulator-integrated tutoring during the first five sessions in addition to the on-screen step-by-step tutorial, which was always available to all participants.[10] The procedure consisted of a complete mastoidectomy with entry into the antrum and a posterior tympanotomy. The first session was longer (60 minutes) and served as an introduction to the procedure and the simulator and in the following 11 procedures there was a strict time limit of 30 minutes for each procedure.

Outcomes and statistics

The simulator auto-saved final-products at 10, 20, and 30 minutes. The stored 30-minute final-product performances were loaded in the simulator in full 3D and analyzed by two blinded expert raters[10] using a 26-item modified Welling Scale for mastoidectomy performance assessment.[24] Each item of the assessment tool is rated binarily with 0 points for an inadequate/incomplete performance and 1 point for adequate/complete performance.

For this study, we extracted the average rating given by the two raters for each of the 26 items rated in every performance (a total of >400 performances). We calculated the mean score for each item for a) sessions 2–12 to establish a general mean, b) session 2–5 to isolate the effect of tutoring regardless of practice condition, c) sessions 2–12 according to whether the massed or distributed practice groups had received initial simulator-integrated tutoring, d) or not received simulator-integrated tutoring, and e) procedures terminated before the allowed time is up or using allowed time (see below).

Further, to map the progression of the procedure during each session, we calculated the volume difference between two save-files using a small computer program developed for this particular purpose (see Figure 1). For each participant and session, we subtracted the auto-save files from 30 and 20 minutes to calculate the amount of bone drilled by the participant during the last 10 minutes of the procedure. In few cases, one or the other save file was corrupt and could therefore not be analyzed and was excluded from study. In addition to calculating the total volume drilled between time points, we also calculated how much of the volume was inside and outside of the reference volume as defined by the built-in simulator-integrated tutor-function.

We defined a procedure as terminated before the time was up if <5 % of the total volume and <2.5 % of the inside reference volume was drilled during the last 10 minutes of the procedure. In this analysis, we only considered the later sessions (session 6–12) because these were non-tutored for all participants.

Statistical analyses were conducted using SPSS (SPSS Inc., Chicago, IL) version 23 for MacOS X with paired samples t-test, chi-squared test of association, and Analysis of Variances (ANOVA).

Ethics

The ethics committee for the Capital Region of Denmark deemed this study to be exempt (H-4-2013-FSP-088).

RESULTS

We identified several parts of the procedure where the novices in general had difficulty achieving a reasonable performance in the VR simulator (mean item score of <0.25 points, Table 1a). These items were primarily related to drilling no holes (in the sigmoid sinus, tegmen, and the external auditory canal) and leaving structures intact (the lateral semicircular canal, the ossicles, and the facial nerve). The simulator-integrated tutor-function significantly increased performance on many of these items compared with the corresponding sessions of non-tutored participants (Table 1b). In addition, the tutor-function increased proper exposure of tegmen tympani and the digastric ridge but also discouraged adequate exposure of the chorda tympani. The effects of the tutor-function did not exceed the effect of repeated, non-tutored practice when comparing all sessions. In relation to repeated practice, time distribution of sessions significantly improved performance on many items including some of the items that in general were of most difficulty for the novices (Table 1c–d). Nonetheless, massed practice where all repetitions were in immediate succession promoted defining the mastoidectomy at the posterior canal wall, adequate thinning of the external auditory canal and leaving the lateral semicircular canal intact when compared with distributed practice.

Overall, we found that in the later sessions (sessions 6–12) 19.3 % of the procedures were terminated at least 10 minutes early, not making full use of the 30 minutes allowed to complete the procedure. The procedures terminated early were not scored significantly higher or lower (mean 14.1 points) than the rest of the procedures (mean 14.3 points)(Table 2). This was also reflected in procedures terminated early and procedures where all time was used being rated equally on most items of the final-product assessment tool (Table 1e). Practice organization, initial simulator-integrated tutoring and session number did not significantly impact on the mean final-product score or the distribution between sessions terminated early or not (Table 2). Four participants (10 %) terminated procedures early three times or more in sessions 6–12 and 11 participants (28 %) always used all the allowed time. Neither of these groups performed significantly different from the other participants.

DISCUSSION

In this study on the repeated performance of novices in VR simulation training of mastoidectomy, we identified several key components of particular difficulty: avoiding drilling holes in the anatomical boundaries of a complete mastoidectomy and violating vital structures. The simulator-integrated tutor-function significantly improved performance in these areas but the effect did often not exceed the effect of repeated practice without simulator-integrated tutoring. Distributed practice had a positive effect on many aspects of the procedure compared with massed practice, including several of the items found to be most difficult. In approximately 1 of 5 procedures, participants did not make full use of the allowed time. However, these performances did not constitute the worst or the best performances. In addition, participants who consistently did not make use of the time or always used all of the time performed equally to the remaining participants. This indicates poor self-assessment skills and a lack of knowledge on when to stop or how to excel.

First of all, we only considered final-product performance and not technical skills or process, which are considered other very important aspects of surgical competency.[24-27] Next, a weakness of this study is the definition of the procedures that were terminated early: the simulator auto-saved the mastoidectomy with 10-minute intervals rather than in smaller intervals, which more precisely could have defined the two different groups in this study. Consequently, we had to define a cut-off in relation to the volume drilled to separate the two groups. Even though a few participants drilled nothing during the last 10 minutes, based on observations, several participants stopped drilling during the last 5–7 minutes. We therefore included an amount of volume based on an analysis of the volume progression at 10, 20, and 30 minutes. Arguably, some of the performances in the group that used all the time had not reached the maximal potential. However, both the early plateau of the average learning curve and the similar score of the participants who always used all the allowed time suggest that generally other factors contribute to the performance ceiling. Overall, fixing the time allowed for the procedure made insights into the self-assessment skills of novices possible. Lastly, the number of participants included was defined by the original study and not based on sample-size calculation for secondary outcomes such as the items analysis in this study. This increases the potential risk of introducing a type II error. However, in each comparison group, a large number of performances were included because each participant had completed 12 sessions of repeated practice. This allowed us to investigate the effect of multiple factors of importance for novice performance—feedback and tutoring, repeated practice, and organization of practice—with a substantial number of performances in each comparison group (between 70 and 230 performances).

A crucial step in improving directed, self-regulated VR simulation training is to map the progression of performance and analyze causes for the learning curve plateau. Another study has similarly identified a plateau in the mastoidectomy performance of novices in a VR simulator[28]: after 4–5 repetitions and measured by a simulator-generated score, cortical mastoidectomy and exposure of the sigmoid sinus did not further improve. The amount of reference volume removed did not improve with repetition in contrast to the time to completion that improved substantially. The more complex task of exposing the incus required more practice and did not plateau during the 6 repetitions studied. Likewise, more difficult parts of the mastoidectomy required markedly more training in the OR for novice residents to gain competency.[12] A similar pattern was found in our data: defining the cortical mastoidectomy and entering the antrum were accomplished in most performances whereas the facial recess was less likely to be adequately exposed. Nonetheless, as reflected in the plateau of the overall learning curve, little improvement occurred with repeated practice. Instead, some items were just difficult to perform satisfactorily for many participants regardless of repeated practice. This might be attributed to the technical fidelity of the simulator, limiting the performance and introducing a ceiling effect. However, as previously reported a few performances reached a very high score and almost one third of the performances were at or above the level of the plateau.[10] This suggests that there are other explanations for the plateau and these issues should be considered in order to improve the performance of novices beyond the initial plateau as discussed in the following.

The built-in tutor-function provides a visual aid, guiding each step of the procedure by highlighting the volume to be drilled in correspondence with the step-by-step tutorial (concurrent feedback). This significantly improved performance on many of the key items of difficulty and suggests that simulator-integrated tutoring has a potential role in directed, self-regulated learning of mastoidectomy. However, many items were in general scored low and the tutor-function could be improved further in combination with the step-by-step guide to better illustrate and explain these areas of difficulty.

A meta-analysis on the effect of feedback in simulation-based procedural skills training in medical education found a moderate but favorable effect on skill outcomes.[17] Concurrent feedback is suggested to be beneficial mainly in complex tasks but the drawback of concurrent feedback is the potential for over-reliance and a performance decline when feedback is ceased.[17] Our findings corroborate this and simulator-integrated tutoring seemed mostly to have a positive effective while on rather than an effect on longer term performance. This emphasizes one of the challenges of tutoring and should also be considered in the context of novice training. Simulator-integrated tutoring should therefore be used to accelerate the initial learning curve and could also be used once in a while for reinforcement but should not be continuously applied.

Distributed practice has consistently been demonstrated to be superior to massed practice[10,29,30] because time allows the consolidation of skills to occur.[31] In general, distributed practice in combination with tutoring had a strong positive effect but interestingly, massed practice in combination with tutoring caused a higher frequency of leaving the lateral semicircular canal intact. This could suggest that the tutor-function entices risky behavior that is modified by immediate repetition where the participants know not to do everything as the tutor-function seem to suggest. This is an excellent example of self-regulation but highlights an issue that needs to be addressed in future improvement of the VR simulator and the integrated tutor function.

As hypothesized, we found that our novices had poor self-assessment skills in relation to the mastoidectomy procedure because many procedures were terminated early while at the same time not being better than the remaining procedures and the additional time would have permitted a better performance if participants had not stopped early. In contrast, other studies have found novices to be able to accurately self-assess surgical technical skills.[9,15]. A possible explanation could relate to mastoidectomy being a more complex and compound procedure than the technical skills studied in other reports. Concerning the learning curve plateauing early, a study on basic knot tying demonstrated that additional practice beyond self-assessed proficiency did not lead to improved skills.[16] The authors suggest an explanation might be that cognitive effort towards further skill acquisition ceased once the novice perceived to have reached proficiency. Cognitive effort is of importance to learning and arrested development occurs when practice is not deliberate.[32] Self-assessment is a complex cognitive skill and the concept of self-assessment constitutes the domains of reflection, self-monitoring and self-directed assessment seeking.[33] Self-assessment skills are vital to effective directed, self-regulated learning.[13] If structured correctly and using precise criteria for self-assessment, a self-guided curriculum can be effective for learning technical surgical skills and to make best use of resources should precede more advanced training with faculty.[9] This emphasizes that a directed, self-regulated technical skills curriculum should also include explicit and specific process goals and the necessary supportive and directive instructions in self-guided learning.[20] It is a responsibility for educators to deliberately design instructions to support directed, self-regulated learning.[18]

All in all, there is much room for improvement of directed, self-regulated VR simulation training and a holistic approach addressing several issues is necessary to overcome the initial plateau in novice performance. Clearly, the instructions could be improved in relation to specific items and emphasize the key elements of difficulty identified. Simulator-integrated tutoring can visually aid some aspects of the procedure but must be used in a way that supports long term learning and adjusted in relation to some items. Self-assessment must be specifically addressed to encourage continuous progression and deliberate practice to avoid arrested development of skills. Supplemental material and instructions could be introduced to accomplish this including introducing participants to what constitutes a good performance. There could also be a role for future developments such as automated formative and summative feedback if implemented correctly within the directed, self-regulated learning framework. Finally, including a range of different of temporal bone models in the VR simulator to reflect the clinical variability in the anatomy of the temporal bone would not only be of importance to mastoidectomy skills training but could also be hypothesized to modify the learning curve plateau by introducing different learning experiences.

CONCLUSIONS

Directed, self-regulated VR simulation training of surgical skills needs a strong instructional design to improve performance beyond an initial plateau and consider not only technical skills but also cognitive aspects such as self-assessment. In VR simulation training of mastoidectomy, novices had difficulty not drilling holes in the outer anatomical boundaries and leaving vital structures intact. Simulator-integrated tutoring improved performance on the aspects but the effect was not sustained when tutoring ceased. Such over-reliance on tutoring emphasizes the need to use tutoring in combination with a more explicit instructional approach and process goals. Novices seem to have poor self-assessment skills in relation to the mastoidectomy procedure and have difficulty in knowing when to stop or how to excel. Deliberate practice and cognitive effort are needed to avoid arrested development and should be considered when designing directed, self-regulated training programs.

REFERENCES

1. Wiet GJ, Stredney D, Wan D. Training and Simulation in Otolaryngology.

Otolaryngol Clin North Am. 2011; 44:1333–1350.
2. George AP, De R. Review of temporal bone dissection teaching: how it was, is and will be. J Laryngol Otol. 2009; 124:119–125.
3. Morris D, Sewell C, Barbagli F, Salisbury K, Blevins NH, Girod S. Visuohaptic simulation of bone surgery for training and evaluation. IEEE Comput Grap Appl. 2006; 26:48–57.
4. Zirkle M, Roberson DW, Leuwer R, Dubrowski A. Using a Virtual Reality Temporal Bone Simulator to Assess Otolaryngology Trainees. Laryngoscope. 2007;117(2):258-263.
5. O'Leary SJ, Hutchins MA, Stevenson DR, et al. Validation of a Networked Virtual Reality Simulation of Temporal Bone Surgery. Laryngoscope 2008;118(6):1040-1046.
6. Wiet GJ, Stredney D, Kerwin T, et al. Virtual temporal bone dissection system: OSU virtual temporal bone system. Laryngoscope 2012; 122:S1–S12.
7. Sørensen 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.
8. Andersen SAW, Foghsgaard S, Konge L, Caye-Thomasen P, Sørensen MS. The effect of self-directed virtual reality simulation on dissection training performance in mastoidectomy. Laryngoscope October 9 2015. [Epub ahead of print].
9. Wright AS, McKenzie J, Tsigonis A, et al. A structured self-directed basic skills curriculum results in improved technical performance in the absence of expert faculty teaching. Surgery 2012; 151:808–814.
10. Andersen SA, Konge L, Caye-Thomasen P, Sørensen MS. Learning Curves of Virtual Mastoidectomy in Distributed and Massed Practice. JAMA Otolaryngol Head Neck Surg 2015; 141:913-918.
11. West N, Konge L, Cayé-Thomasen P, Sørensen MS, Andersen SAW. Peak and ceiling effects in final-product analysis of mastoidectomy performance. J Laryngol Otol 2015; 129:1091–1096.
12. Francis HW, Masood H, Laeeq K, Bhatti NI. Defining milestones toward competency in mastoidectomy using a skills assessment paradigm. Laryngoscope 2010; 120:1417–1421.
13. Brydges R, Dubrowski A, Regehr G. A New Concept of Unsupervised Learning:

Directed Self-Guided Learning in the Health Professions. Acad Med 2010; 85:S49–S55.
14. Mcgaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ. 2010; 44:50–63.
15. MacDonald J, Williams RG, Rogers DA. Self-assessment in simulation-based surgical skills training. Am J of Surg 2003; 185:319–322.
16. Jowett N, LeBlanc V, Xeroulis G, MacRae H, Dubrowski A. Surgical skill acquisition with self-directed practice using computer-based video training. Am J Surg 2007; 193:237–242.
17. Hatala R, Cook DA, Zendejas B, Hamstra SJ, Brydges R. Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. Adv in Health Sci Educ 2013; 19:251–272.
18. Brydges R, Manzone J, Shanks D, et al. Self-regulated learning in simulation-based training: a systematic review and meta-analysis. Med Educ 2015; 49:368–378.
19. Butler DL, Brydges R. Learning in the health professions: what does self-regulation have to do with it? Med Educ 2013; 47:1057–1059.
20. Brydges R, Carnahan H, Safir O, Dubrowski A. How effective is self-guided learning of clinical technical skills? It's all about process. Med Educ 2009; 43:507–515.
21. Sørensen MS, Dobrzeniecki AB, Larsen P. The visible ear: a digital image library of the temporal bone ORL J Otorhinolaryngol Relat Spec 2001; 64:378–381.
22. The Visible Ear Simulator. http://ves.alexandra.dk. Accessed 19 February 2016.
23. Andersen SAW, Mikkelsen PT, Konge L, Caye-Thomasen P, Sørensen MS. Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation. Laryngoscope 2016; 126:E74-9.
24. Andersen SAW, Caye-Thomasen P, Sorensen MS. Mastoidectomy performance assessment of virtual simulation training using final-product analysis. Laryngoscope 2014; 125:431–435.
25. Zirkle M, Taplin MA, Anthony R, Dubrowski A. Objective Assessment of Temporal Bone Drilling Skills. Ann Otol Rhinol Laryngol 2007; 116:793–798.
26. Laeeq K, Bhatti NI, Carey JP, et al. Pilot testing of an assessment tool for competency in mastoidectomy. Laryngoscope 2009; 119:2402–2410.

27. Francis HW, Masood H, Chaudhry KN, et al. Objective Assessment of Mastoidectomy Skills in the Operating Room. Otol Neurotol 2010l; 31:759–65.
28. Nash R, Sykes R, Majithia A, Arora A, Singh A, Khemani S. Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator. J Laryngol Otol 2012; 126:663–669.
29. Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training. Surg Endosc 2002; 16:957–961.
30. Moulton C-AE, Dubrowski A, MacRae H, Graham B, Grober E, Reznick R. Teaching Surgical Skills: What Kind of Practice Makes Perfect? Ann Surg 2006; 244:400–9.
31. Shea CH, Lai Q, Black C, Park J-H. Spacing practice sessions across days benefits the learning of motor skills. Human Movement Science 2000; 19:737–760.
32. Ericsson KA. Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. Acad Med 2004; 79:S70–S81.
33. Eva KW, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof 2008; 28:14–19.

Table I - Mean mastoidectomy final-product scores	a) All		b) Both practice conditions			c) Tutored group			d) Non-tutored group			e) Using allowed time		
	Mean	Maximum§	Tutored sessions (2–5)	Non-tutored sessions (2–5)	Significance of difference	Massed practice	Distributed practice	Significance of difference	Massed practice	Distributed practice	Significance of difference	Terminates procedure before time is up (session 6–12)	Uses all time allowed (session 6–12)	Significance of difference
Mastoidectomy margins defined at														
1. Temporal line	0.99	1.00	0.99	0.97	ns	0.99	0.99	ns	0.99	0.99	ns	1.00	0.99	ns
2. Posterior canal wall	0.88	0.97	0.92	0.97	ns	0.99	0.76	p<0.001*	0.93	0.89	ns	0.89	0.84	ns
3. Sigmoid sinus	0.98	1.00	0.97	0.99	ns	0.97	0.99	ns	0.98	0.98	ns	0.98	0.98	ns
Antrum mastoideum														
4. Antrum entered	0.99	1.00	0.99	0.99	ns	0.99	1.00	ns	1.00	1.00	ns	0.99	1.00	ns
5. Lateral semicircular canal exposed	0.98	1.00	0.97	0.99	ns	0.98	0.98	ns	0.99	0.98	ns	0.96	0.99	p<0.05
6. Lateral semicircular canal intact	0.16	0.35	0.36	0.10	p<0.001*	0.32	0.10	p<0.001*	0.13	0.13	ns	0.17	0.11	ns
Sigmoid sinus														
7. Exposed, no overhang	0.97	1.00	0.99	0.97	ns	0.94	0.98	ns	0.95	0.99	ns	0.97	0.96	ns
8. No cells remain	0.66	0.76	0.59	0.67	ns	0.44	0.73	p<0.001*	0.67	0.76	ns	0.64	0.68	ns
9. No holes	0.07	0.13	0.15	0.03	p<0.002*	0.06	0.13	ns	0.02	0.08	p<0.003	0.09	0.05	ns
Sinodural angle														
10. Sharp	0.73	0.84	0.66	0.76	ns	0.53	0.78	p<0.001*	0.76	0.80	ns	0.67	0.75	ns
11. No cells remain	0.38	0.53	0.30	0.36	ns	0.23	0.38	p<0.003	0.49	0.41	ns	0.42	0.40	ns
Tegmen mastoideum/tympani														
12. Attic/tegmen tympany exposed	0.80	0.92	0.95	0.78	p<0.001*	0.77	0.86	ns	0.76	0.78	ns	0.72	0.76	ns
13. Ossicles intact (untouched)	0.24	0.37	0.46	0.20	p<0.001*	0.33	0.27	ns	0.19	0.20	ns	0.14	0.21	ns
14. Tegmen mastoideum exposed	0.98	1.00	0.99	0.98	ns	0.95	0.99	ns	1.00	0.98	ns	1.00	0.98	ns
15. No cells remain	0.43	0.53	0.53	0.41	ns	0.33	0.56	p<0.001*	0.41	0.40	ns	0.37	0.43	ns
16. No holes	0.09	0.13	0.08	0.07	ns	0.04	0.13	p<0.001*	0.04	0.13	p<0.001*	0.03	0.10	p<0.05
Mastoid tip														
17. Digastric ridge exposed	0.35	0.49	0.51	0.35	p<0.02	0.30	0.41	ns	0.28	0.37	ns	0.43	0.28	p<0.01
18. Digastric ridge followed towards stylomastoid foramen	0.09	0.16	0.14	0.09	ns	0.06	0.11	ns	0.06	0.12	ns	0.10	0.08	ns
19. No cells remain	0.05	0.10	0.15	0.00	p<0.001*	0.05	0.07	ns	0.03	0.04	ns	0.03	0.04	ns
External auditory canal														
20. Thinning of the posterior canal wall	0.78	0.93	0.85	0.86	ns	0.88	0.68	p<0.001*	0.83	0.76	ns	0.80	0.74	ns
21. No cells remain	0.33	0.47	0.52	0.32	p<0.001*	0.39	0.36	ns	0.25	0.33	ns	0.31	0.28	ns
22. No holes	0.05	0.10	0.09	0.04	ns	0.08	0.05	ns	0.03	0.06	ns	0.05	0.04	ns
Facial nerve														
23. Facial nerve identified (vertical part)	0.94	1.00	0.88	0.94	ns	0.91	0.94	ns	0.92	0.97	ns	0.89	0.96	p<0.05
24. No exposed nerve sheath	0.14	0.26	0.13	0.11	ns	0.15	0.15	ns	0.08	0.19	p<0.01	0.12	0.18	ns
25. Tympanic chorda exposed	0.90	0.96	0.78	0.92	p<0.02	0.83	0.89	ns	0.90	0.95	ns	0.87	0.93	ns
Posterior tympanotomy														
26. Facial recess completely exposed	0.45	0.63	0.40	0.36	ns	0.39	0.58	p<0.001*	0.37	0.43	ns	0.46	0.51	ns
Total mean of all sessions	14.43		15.35	14.23		13.90	14.87		14.05	14.70		14.09	14.27	

* Significant at the level of Bonferroni correction (p=0.002)

§ Maximum mean observed in a session

Table 1. Mean scores for each item of the modified Welling Scale for mastoidectomy final-product analysis: a) the overall mean across all sessions and all groups, b) mean for sessions 2–5 according to tutoring of these sessions regardless of practice condition, c) mean for participants tutored by the simulator-integrated tutor-function according to practice condition, d) mean for non-tutored participants according to practice condition, and e) mean for sessions 6–12 according to whether the participant terminated the procedure before the allowed 30 minutes or used all the time.

	Terminates procedure before time is up		Uses all time allowed		
	n	Mean final-product score (95 % CI)	n	Mean final-product score (95 % CI)	Significance of difference between groups
a. All	52	14.1 (13.3–14.8)	217	14.3 (13.9–14.6)	ns
b. Practice condition					
Massed	30	13.6 (12.7–14.4)	99	13.4 (12.9–13.9)	ns
Distributed	22	14.8 (13.5–16.1)	118	14.8 (14.3–15.3)	ns
c. Initial simulator-integrated tutoring (session 2–5)					
Tutored	20	14.0 (12.8–15.2)	114	13.8 (13.2–14.3)	ns
Non-tutored	32	14.1 (13.1–15.1)	103	14.6 (14.1–15.1)	ns
d. Session					
#6	6	13.8 (11.7–15.8)	34	14.2 (13.5–14.9)	ns
#7	5	12.5 (11.7–12.3)	32	13.3 (11.9–14.7)	ns
#8	7	14.1 (12.0–16.1)	32	14.3 (13.5–15.2)	ns
#9	4	12.9 (9.7–16.1)	33	14.1 (13.2–15.0)	ns
#10	11	14.9 (13.2–16.6)	28	14.8 (13.7–16.0)	ns
#11	8	14.9 (12.5–17.2)	29	13.9 (13.0–14.8)	ns
#12	11	14.1 (11.5–16.7)	29	14.6 (13.6–15.6)	ns

Table 2. Mean final-product scores for sessions in which the participant stopped drilling before the allowed 30 minutes was up and for sessions where all the time was used for a) the overall mean, and according to b) practice condition, c) initial simulator-integrated tutoring in session 1–5, and d) session number.

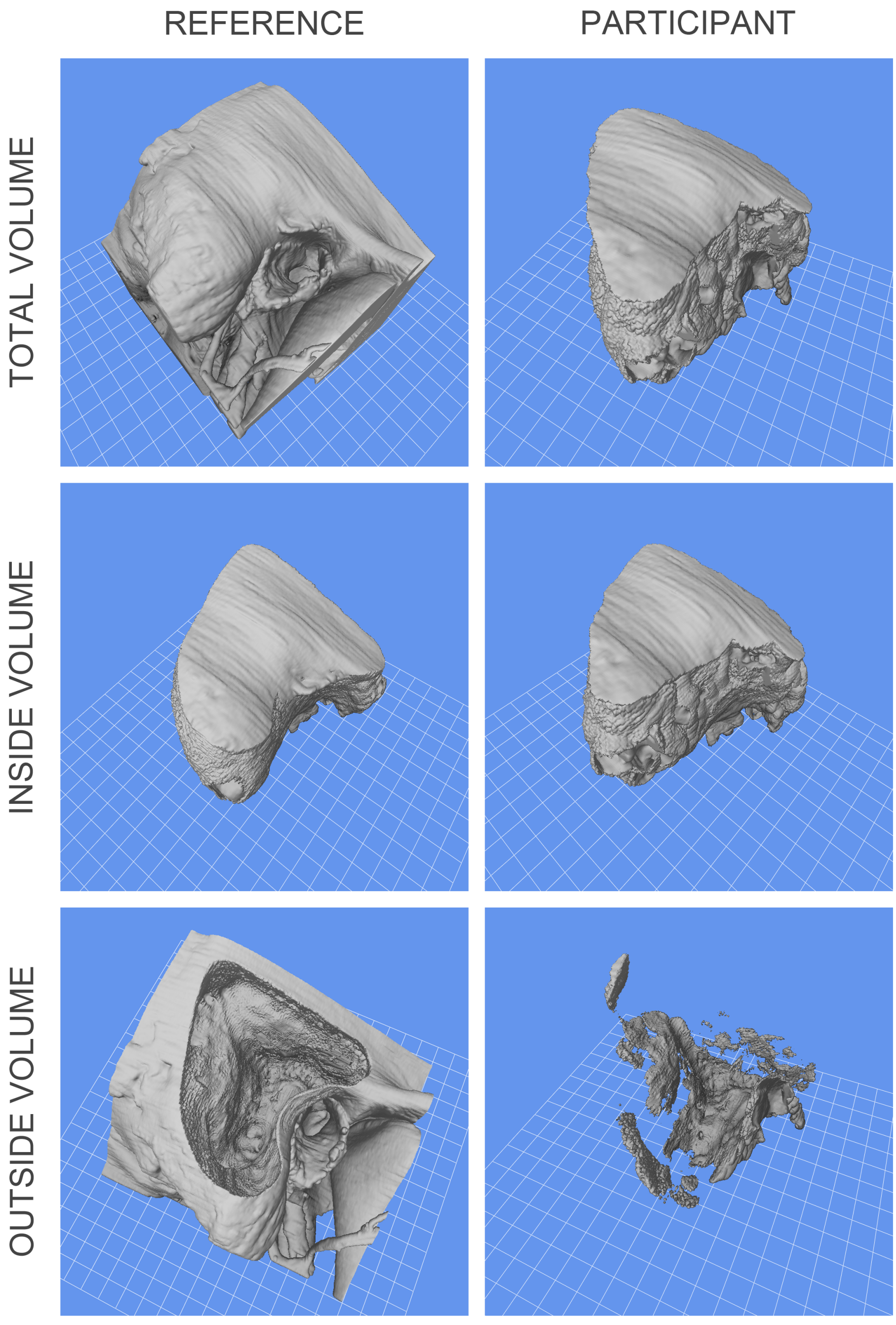

Figure 1. Examples from the volume analysis tool. The volumes (left panel) of the temporal bone in the simulator with the total volume of temporal bone included in the simulator (top), the reference volume for a complete mastoidectomy and posterior tympanotomy (middle) and the volume outside of the reference volume (bottom). An example of a participants drilling (right panel) with the total volume drilled (top), the volume drilled inside the reference volume (middle) and the volume drilled outside the reference volume (bottom).