The Journal of International Advanced Otology

J Int Adv Otol 2023; 19(2): 99-104 • DOI: 10.5152/iao.2023.22753

Original Article

Learning Curves in Directed Self-Regulated Virtual Reality Training of Mastoidectomy and the Role of Repetition and Motivation

Hagar al Fartoussi[1], Mads Sølvsten Sørensen[1,2], Steven Arild Wuyts Andersen[1,2]

[1]Copenhagen Hearing and Balance Center, Department of Otorhinolaryngology, Rigshospitalet, Copenhagen, Denmark
[2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

ORCID IDs of the authors: H.F. 0000-0001-6998-6718, M.S.S. 0000-0003-1146-302X, S.A.W.A. 0000-0002-3491-9790.

Cite this article as: al Fartoussi H, Sørensen MS, Andersen SAW. Learning curves in directed self-regulated virtual reality training of mastoidectomy and the role of repetition and motivation. *J Int Adv Otol*. 2023;19(2):99-104.

BACKGROUND: Mastoidectomy is a complex procedure which can be trained on human cadaveric temporal bones or simulation models. The number of repetitions offered in most training curricula is considerably less than what is normally required for motor skills acquisition in crafts or sports. Directed, self-regulated virtual reality simulation training may provide unlimited repetitions but the effect on learning of extended but unsupervised training is unknown. This study recorded extended learning curves of novices in virtual reality simulation mastoidectomy training.

METHODS: Six medical students used the visible ear temporal bone simulator at home for 100 repetitions. Virtual temporal bones were later assessed by 2 blinded experts on a 26-point modified Welling Scale.

RESULTS: Four participants completed 100 procedures each and 2 participants completed 50 procedures. Learning curves and plots of time demonstrated great variation: one participant improved gradually during the first 50 procedures and sustained a high performance; another participant achieved only 16 points after 50 procedures; a third participant demonstrated mediocre performances between 10 and 15 points but only used about 5 minutes per procedure. The remaining 3 participants achieved high but fluctuating scores with very limited time use per procedure. Their score per time exceeds the performance of experienced otosurgeons and suggests the use of save/restore gaming strategies to inflate their performance.

CONCLUSION: Deliberate learners may reach proficiency in virtual reality simulation of mastoidectomy after 50 repetitions. However, even 100 repetitions cannot guarantee proficiency if motivation fails. Creative "gaming" behavior must be considered and opposed by motivation, supervision, testing, and certification.

KEYWORDS: Mastoidectomy, simulation-based training, learning curves, curriculum development, evidence-based medical education

INTRODUCTION

Surgical education has undergone a paradigm shift in recent decades – from the principle of "see one, do one" toward evidence-based training.[1] This change has been fueled by many factors, including new requirements for training efficacy and patient safety, increasing numbers of trainees, decreasing work hours, and poor availability of donated human materials. Altogether, this necessitates the use of alternative learning platforms.[2] For this to be successful, there is a need to understand how surgical technical skills are learned and to develop modern training methods that support high-quality learning.

Learning curves are essential in understanding skills acquisition and are dependent not only on learner characteristics but also on instructional design and learning context.[3] Learning curves can be used to inform best practice implementation and organization of training. Grantcharov and Funch-Jensen[4] identified 4 learning curve patterns in the acquisition of basic laparoscopic skills: A few learners demonstrate proficiency almost from the beginning; the majority of learners achieve a predefined expert criterion after a certain number of repetitions; another small group of learners do improve some with repetition but are unable to achieve proficiency within the time provided; and finally, a few learners consistently underperform with no true improvement. Structured and reliable assessment of performance is key to measure learning curves in surgical technical skills training: the introduction of

Corresponding author: Steven Arild Wuyts Andersen, e-mail: stevenarild@gmail.com
Received: March 28, 2022 • **Accepted:** August 2, 2022 • **Publication Date:** March 24, 2023
Available online at www.advancedotology.org

Objective Structured Assessment of Technical Skills[5] pioneered such systematic assessment in the early 1990s and inspired the development of specific assessment tools for a number of different surgical skills and procedures.

Virtual reality (VR) simulation training in temporal bone surgery is an example of a newer learning modality that allows otorhinolaryngology trainees to acquire basic mastoidectomy skills in a patient-safe environment, independent of service duties and without the presence of supervisors, supporting individual training needs through directed self-regulated learning. Such VR temporal bone simulation training is supported by evidence of the efficacy and validity of training early novices[6] and several validated tools for structured performance assessment exist.[7] For novices training in a VR temporal bone simulator, performances seemingly plateau at 4-9 performances[8,9] at a level well below that of experienced surgeons, who had a final product mean score of 19.6 points out of 26 points on the modified Welling Scale.[10] Frequently, we find that novices make damage to the facial nerve and inner ear structures, which limit their performance and cause a seeming plateau in their learning curve. Despite problems with ceiling effects that might arise from the simulation or the assessment itself,[11] we also sometimes observe novice performances achieving the maximum score or close to it. None of the current studies on temporal bone simulation include more than 18 repetitions, and the majority of novices might need considerably more practice to achieve a satisfying performance. So far, additional learning supports such as simulator-integrated tutoring[12] or structured self-assessment[13] to improve the quality of the training itself have failed to increase performance beyond the early plateau. It is therefore possible that previous interventions trying to improve learning have been too short and that some learners might actually improve but at a slower rate as reported by Grantcharov and Funch-Jensen.[4] Even for less complex surgical skills such as laparoscopic box training, very few studies include a high number of repetitions,[14] while the repetition of basic procedures in sports, arts, and craftsmanship is generally at a higher order of magnitude.[15] Our research question is therefore: Can substantially more simulation-based training counter the observed learning curve plateau in a directed, self-regulated training program? This pilot study aims to examine the effect on the learning of numerous repetitions in VR simulation training of novices in temporal bone surgery.

METHODS

Virtual Reality Simulation Platform

The Visible Ear Simulator (VES)[16] version 3.5 is a freeware real-time, 3-dimensional (3D), VR temporal bone surgical simulator that can be downloaded from the Internet.[17] The software runs on a standard personal computer with an Nvidia GeForce GTX/RTX™ graphics card. A Geomagic Touch (3D Systems, Rock Hill, SC, USA) haptic device is recommended for intuitive drilling with force feedback.

A built-in instructional guide is provided in a side panel on the simulation screen and resembles a traditional temporal bone manual. Each step of the procedure is explained with brief written instructions and a picture from the simulator illustrating the step with key anatomical landmarks indicated. The integrated tutor-function provides optional color coding of the volume of bone to be drilled in each step corresponding to the built-in guide, and this intuitively visualizes the volume to be removed directly on the interactive model in the central workspace.

Participants and Setting

The first author contributed to the study and further recruited 5 fellow participants (Table 1). They were all medical students from the Faculty of Health and Medical Sciences, University of Copenhagen, Denmark and represented true novices to the procedure. The study took place at Department of Otorhinolaryngology, Rigshospitalet, Copenhagen, Denmark, from December 2020 to April 2021. All participants volunteered and received no compensation or study credits for participation.

Study Design and Intervention

In this non-comparative pilot study, participants performed repeated virtual mastoidectomies in a VR temporal bone surgical simulator.

In order to facilitate the completion of a large number of procedures, the typical setup for drilling under supervision in the department was replaced by at-home drilling using high-performance laptop PCs equipped with a Geomagic Touch haptic device, which were loaned to the participants.

Each participant received a brief, standardized instruction after which they were asked to perform 100 repetitions (i.e., identical anatomical mastoidectomy procedures) in the simulator, distributed over the study period and at their own pace. The participants could use simulator-integrated tutoring by color coding during the first procedure only and were instructed to disable this function for the rest of the procedures. Participants had access to step-by-step on-screen instructions for the procedure at all times and they could contact the study investigators if they experienced any technical problems during the training at home. At the end of each repetition, the virtually drilled temporal bone model was saved manually on the PC by the trainee together with the exercise number. The duration of the procedure was logged by the simulator in the save file.

Sample Size

The sample size was one of convenience and limited by the extensive time required by study participants. No financial compensation or performance-related incentives were offered due to concerns about bias and it was not possible to recruit further participants within the study period.

Outcomes and Statistics

The saved final product of every fifth exercise was loaded in the simulator and assessed independently by 2 experienced raters (M.S.S.,

Table 1. Participant Background Data

Participant	Age	Sex	Gaming Experience	Interest in Surgery
A	22	Female	None	High
B	21	Female	10 h/week	High
C	22	Female	None	Low
D	23	Female	None	High
E	27	Female	None	Very high
F	19	Male	15 h/week	Low

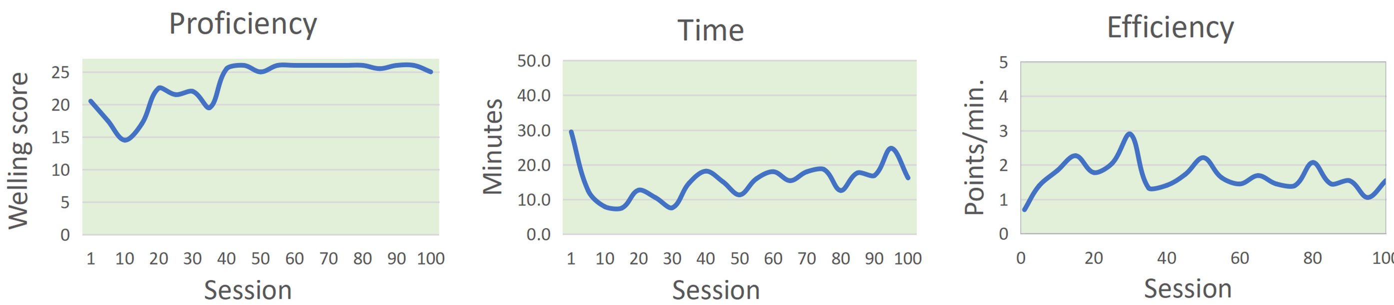


Figure 1. Proficiency, time, and efficiency plots for participant A.

S.A.W.A.) blinded to participant ID and performance number. They used the 26-item modified Welling Scale, where each item is rated binarily with 0 points for an inadequate/incomplete performance and 1 point for adequate/complete performance.[18] Individual learning curves were produced for each participant based on this final-product assessment along with time recorded by the simulator. Statistical analysis was performed using standard descriptive statistics for the demographic data, and learning curves were plotted in Microsoft Excel version 16.48 (Redmond, WA, USA).

Ethics

The study was deemed exempt by the ethics committee for the the Capital Region of Denmark (H-20078583).

RESULTS

Participant Characteristics

Six participants were recruited: 5 were female; the mean age of the participants was 22.6 years; and the median semester of the study was 6th. Two participants (D and F) had a leisure computer gaming background whereas the remaining 4 participants did not engage in computer gaming activities (Table 1). Four participants (A, B, C, and D) completed all 100 procedures, whereas the remaining 2 (E and F) completed only 50 procedures within the study period.

Individual Performance Curves

A mean learning curve of average performance as a function of repetition number across all participants proved irrelevant due to a large, unexpected variance among the 6 observed performance patterns. The observed learning curve patterns are classified according to Grantcharov and Funch-Jensen's terminology in Table 1.

Participant A (Figure 1) achieved a tutored first score of 20.5 points after which the performance without tutoring (i.e., without color coding) dropped to 14 points for 10 repetitions, then gradually increased to reach a steady high of 25-26 points after 45-50 repetitions. The time consumption of participant A displayed an initial decrease through 10 repetitions, followed by a slight gradual increase toward 20 minutes maintained into the high-performance plateau. Efficiency measured as Welling Scale points per minute of drilling was moderate and constant during the study.

Participant E (Figure 2) only completed 50 repetitions and the performance score increased from 10 to 16 points with generally low time use, decreasing from 15 to less than 10 minutes. The efficiency increased gradually from 1 to more than 3 points per minute.

Participant C (Figure 3) showed a slight initial improvement and then plateaued on a level below 15 points. The corresponding time consumption was very low and decreased early to less than 5 minutes per session. Efficiency showed no significant development.

Participant D (Figure 4) obtained a high score right from the start after drilling for 100 minutes. Without tutoring, the time use dropped rapidly to a very low level around 10 minutes, but the performance stayed high but varied around 20 points. Efficiency showed a reciprocal pattern with an initial low followed by a moderate to high constant level.

Participants B and F (Supplementary Figure 1) completed 100 and 50 sessions, respectively, and presented a performance pattern similar to participant D with a relatively high performance but using very little time for each procedure.

DISCUSSION

The present data could not support the extraction of a universal mean learning curve because of large variability between the individual curves. The only general patterns in the results were: 1) the temporary performance drop as tutorial color coding was withdrawn,

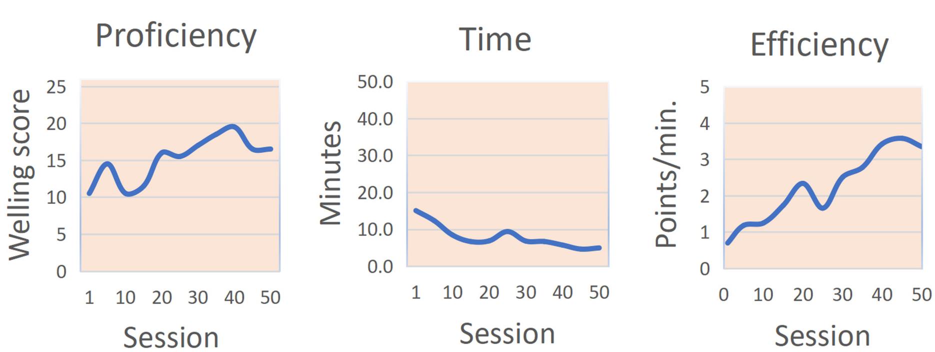


Figure 2. Proficiency, time, and efficiency plots for participant E.

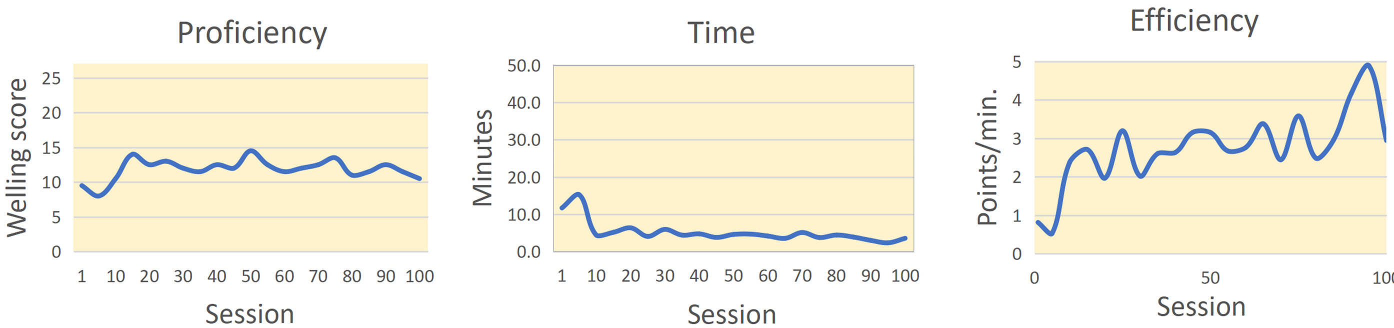


Figure 3. Proficiency, time, and efficiency plots for participant C.

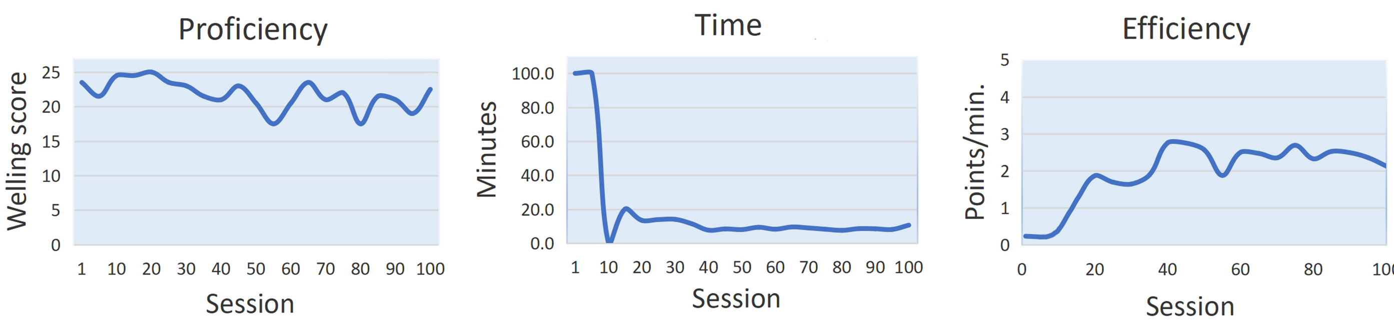


Figure 4. Proficiency, time, and efficiency plots for participant D.

and, 2) the gradually increasing points per minute efficiency, which was found even in cases where the final-product performance did not improve sufficiently.

Increasing efficiency may somehow reflect the learning of mastoidectomy skills, but for patient safety reasons, increasing the final-product rating seems to be a more important training goal. The latter was demonstrated in the learning curve of participant A, who reached a stable level of proficiency after 50 repetitions but this did not come with an increased efficiency.

The performances of some of the participants fit the classification by Grantcharov and Funch-Jensen (Table 2). After 50 repetitions, participant A reached a stable performance level at maximum score (corresponding to "group 2") and the final product was a full resection of the mastoid air space with no harmful exposure of soft tissue limiting structures such as nerves, major vessels, or inner ear spaces. This suggests that the early plateau observed in previous studies[9,19] in some cases might be overcome simply by further repetition—a training principle already well-proven for other psychomotor skills such as playing a musical instrument. Most likely, this participant exhibited traits associated with deliberate practice, that is, purposefully refining skills with cognitive engagement in the task.[15] This also suggests that the amount of temporal bone training offered to most otorhinolaryngology residents as in-house or temporal bone course training should be increased to this level for candidates training for temporal bone surgery to ensure adequate competencies before supervised surgery on patients.

Participants E and C matched the Grantcharov and Funch-Jensen groups 3 and 4: participant E improved to some extent but without ever reaching proficiency at 50 procedures and participant C showed no obvious improvement throughout 100 procedures. A shared characteristic of participants E and C was their very low time use on each procedure. The choice to complete a full mastoidectomy as a novice using only 5-10 minutes may reflect lack of motivation, poor self-assessment, or both, which appears to obstruct learning regardless of number of repetitions. These participants might need further support such as instant feedback provided by the simulator to motivate, guide, and inform the trainee of their progress during the drilling activity. The use of summative feedback at the end of each session seems to motivate the learner and improve self-assessment.[20] Such integrated learning supports are a feature unique to VR simulation as this cannot easily be integrated in training on physical models. Further, VR simulation might lend itself to gamification strategies adopted from leisure PC gaming, which may prove further useful in

Table 2. Participants A–F Classified by Grantcharov and Funch-Jensen's Learning Curve Patterns

Learning Curve Pattern	Participant
Group 1: Demonstrates proficiency from the beginning of the learning session.	None
Group 2: Advances with practice and achieve predefined expert criteria.	A
Group 3: Demonstrates slow and steady skills acquisition but is unable to achieve proficiency level.	E
Group 4: Underperforms from the beginning and demonstrates no skills improvement.	C
Group 5 (new, this study): Achieves a high performance score using very little time. Incompatible with novice level, suggesting the use of gaming strategies.	B, D, F

future training routines for directed self-regulated learning in temporal bone surgery.

None of the participants demonstrated proficiency from the beginning (Grantcharov and Funch-Jensen's group 1) most likely because of the high complexity of the mastoidectomy task and the extensiveness of the procedure in contrast to the simple laparoscopic training tasks studied by Grantcharov and Funch-Jensen. The ability of some learners to achieve proficiency faster than others is most likely multifactorial and might be related to for example innate ability, motivation, ability for learning, cognitive processes, and gaming experience.

Interestingly, our participants B, D, and F displayed some characteristics not accounted for in the Grantcharov and Funch-Jensen learner classification. These participants had unexpectedly high and variable performance scores from the start of training and throughout the entire study period despite using extremely little time on each procedure. Their results are unique when compared to novice performances measured previously during our controlled studies where participants were trained at the simulation center or department.[9,12,13,19] Further, this very short amount of time to achieve such an excellent performance can only occasionally be reproduced in the simulator by expert surgeons. It is therefore possible that the results were enhanced through the unforeseen use of alternative "gaming" strategies at home such as sustained color tutoring, using the "undo drilling" function to rewind drilling and thereby repair damage made to structures, and/or the "save scenario" and "load scenario" commands to restore a well-drilled previously saved performance as a shortcut. These simulator functions were, except for the color tutoring and save function, not introduced to the participants. However, they could have been accessed in the menus and help function by the participants if they explored these. Indeed, D and F both reported leisure gaming skills on their inclusion questionnaire. This suggests that a proxy goal of achieving a high score by the learners may possibly have overruled our intended pursuit of true proficiency by meticulous drilling and sustained repetition. Participants were not aware that the drilling time was recorded in their save files. In other words, control mechanisms are needed, and this should be considered in the strategies of directed, self-regulated learning when trainees are asked to achieve a specific objective such as performing a defined number of procedures.

Overall, our study has important implications for the organization of self-directed training.

Self-regulated VR simulation training at home is convenient especially when many repetitions are needed. However, it is crucial to consider that learning curves are individual and that some trainees will consistently demonstrate proficiency before others. As motivation plays a key role in learning and will need to be high, especially in the context of many repetitions and training at home, addressing motivation, providing learning support and relevant feedback cannot be underestimated. Further, the present results demonstrate that the sheer number of documented sessions completed is not enough proof of proficiency. At the end of the training period, the trainee should perform a few supervised sessions for rating and certification. Moreover, we found that there is a need for training in an order of magnitude that is considerably larger than what is offered at many training institutions.[2] The amount of training offered to most otorhinolaryngology residents as in-house or temporal bone course training should be increased and for candidates training for temporal bone surgery, >50 mastoidectomy repetitions are needed to ensure that the learning potential of simulation-based training is exhausted. Providing opportunity to practice more than 50 mastoidectomy procedures would require the otosurgical curriculum to supplement traditional practice on human cadaveric temporal bone specimens with additional training on inexpensive and convenient models such as VR simulation and 3D-printed temporal bone models.

The main strength of this pilot study is the high number of repetitions compared with previous studies. This adds important new information, namely about the impact of *quantitative* aspects of temporal bone surgical training (i.e., number of repetitions) and defining the significance of the number of repetitions when training mastoidectomy. The main limitation is that we only included few participants, which is explained by the difficulty of recruiting volunteers for such time-consuming participation, and that they were medical student novices rather than otorhinolaryngology residents. Also, we only considered final-product performance and no other important aspects of technical skills performance such as process. Finally, participants were not provided with feedback or scores during their training, which could also reduce their motivation and ultimately performance.

CONCLUSION

In this pilot study, we found that high-repetition practice (>50 repetitions) might be beneficial in overcoming the learning curve plateau for some learners, whereas others do not show any progress most likely because of lack of motivation. We found learning curve patterns that corroborate 3 out of 4 types observed by Grantcharov and Funch-Jensen's and also a new fifth group. The latter consisted of novice learners who used very little time while achieving high-performance scores, most likely through the use of gaming techniques in the VR simulation environment. The importance of deliberate practice cannot be stressed enough as continued cognitive engagement in the learning task is paramount to refine skills and achieve an excellent performance. Altogether, this study adds that control mechanisms need to be included in prolonged self-directed training programs to support learning and continued progress. This important in the design of simulation-based training and certification of surgeons.

Ethics Committee Approval: Ethical committee approval was received from the Ethics Committee of the Capital Region of Denmark (H-20078583).

Informed Consent: Informed consent was obtained from each patient included in the study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – M.S.S.; Design – H.F., S.A.; Supervision – M.S.S., S.A.; Materials – M.S.S.; Data Collection and/or Processing – H.F.; Analysis and/or Interpretation – H.F., S.A., M.S.S.; Literature Review – H.F., S.A.; Writing – H.F., M.S.S.; Critical Review – S.A.

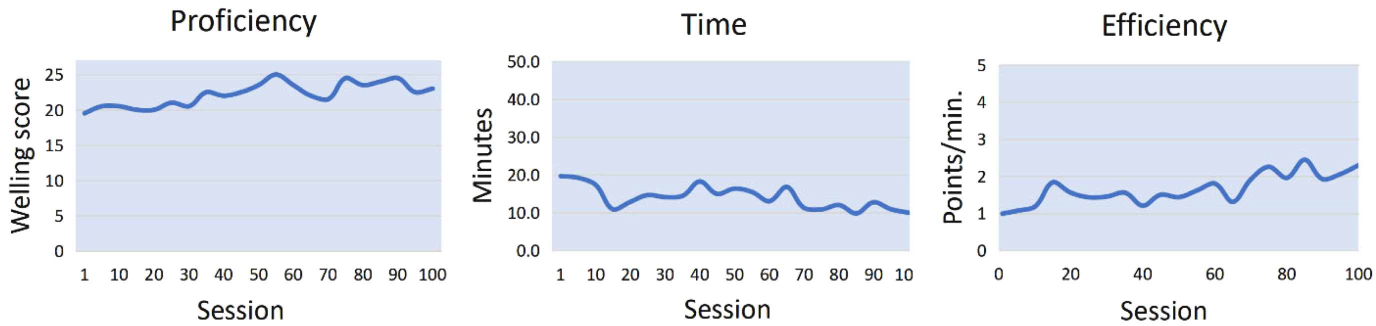
Declaration of Interests: The authors declare that they have no conflict of interest.

Funding: The authors declared that this study has received no financial support.

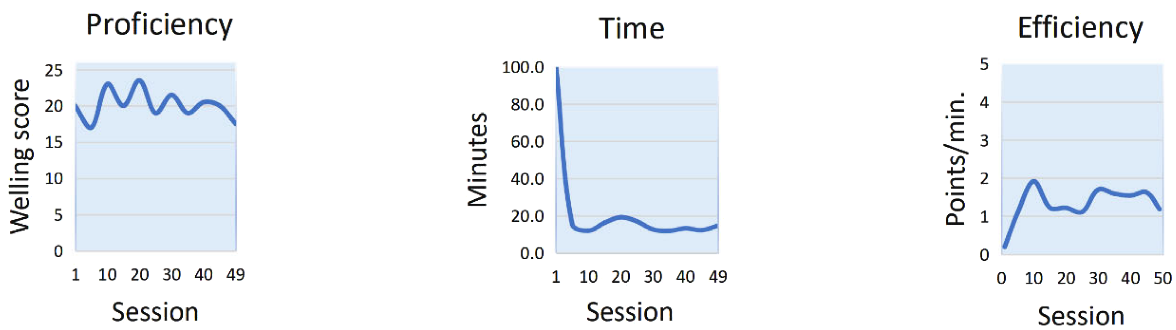
REFERENCES

1. Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. *N Engl J Med*. 2006;355(25):2664-2669. [CrossRef]
2. Frithioff A, Sørensen MS, Andersen SAW. European status on temporal bone training: a questionnaire study. *Eur Arch Otorhinolaryngol*. 2018;275(2):357-363. [CrossRef]
3. Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med*. 2015;90(8):1034-1042. [CrossRef]
4. Grantcharov TP, Funch-Jensen P. Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition. *Am J Surg*. 2009;197(4):447-449. [CrossRef]
5. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84(2):273-278. [CrossRef]
6. Lui JT, Hoy MY. Evaluating the effect of virtual reality temporal bone simulation on mastoidectomy performance: A meta-analysis. *Otolaryngol Head Neck Surg*. 2017;156(6):1018-1024. [CrossRef]
7. Sethia R, Kerwin TF, Wiet GJ. Performance assessment for mastoidectomy. *Otolaryngol Head Neck Surg*. 2017;156(1):61-69. [CrossRef]
8. 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(7):663-669. [CrossRef]
9. Andersen SA, Konge L, Cayé-Thomasen P, Sørensen MS. Learning curves of virtual mastoidectomy in distributed and massed practice. *JAMA Otolaryngol Head Neck Surg*. 2015;141(10):913-918. [CrossRef]
10. Andersen SAW, Mikkelsen PT, Sørensen MS. Expert sampling of VR simulator metrics for automated assessment of mastoidectomy performance. *Laryngoscope*. 2019;129(9):2170-2177. [CrossRef]
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(11):1091-1096. [CrossRef]
12. Andersen SAW, Mikkelsen PT, Sørensen MS. The effect of simulator-integrated tutoring for guidance in virtual reality simulation training. *Simul Healthc*. 2020;15(3):147-153. [CrossRef]
13. Andersen SAW, Guldager M, Mikkelsen PT, Sørensen MS. The effect of structured self-assessment in virtual reality simulation training of mastoidectomy. *Eur Arch Otorhinolaryngol*. 2019;276(12):3345-3352. [CrossRef]
14. Louridas M, Szasz P, Fecso AB, et al. Practice does not always make perfect: need for selection curricula in modern surgical training. *Surg Endosc*. 2017;31(9):3718-3727. [CrossRef]
15. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med*. 2004;79(Supplement):S70-S81. [CrossRef]
16. Trier P, Noe KØ, Sørensen MS, Mosegaard J. The visible ear surgery simulator. *Stud Health Technol Inform*. 2008;132:523-525.
17. *The Visible Ear Simulator*. https://ves.alexandra.dk Accessed 11 Feb 2022.
18. Andersen SA, Cayé-Thomasen P, Sørensen MS. Mastoidectomy performance assessment of virtual simulation training using final-product analysis. *Laryngoscope*. 2015;125(2):431-435. [CrossRef]
19. Andersen SAW, Foghsgaard S, Cayé-Thomasen P, Sørensen MS. The effect of a distributed virtual reality simulation training program on dissection mastoidectomy performance. *Otol Neurotol*. 2018;39(10):1277-1284. [CrossRef]
20. Frithioff A, Frendø M, von Buchwald JH, Trier Mikkelsen P, Sølvsten Sørensen M, Arild Wuyts Andersen S. Automated summative feedback improves performance and retention in simulation training of mastoidectomy: a randomised controlled trial. *J Laryngol Otol*. 2022;136(1):29-36. [CrossRef]

SDC1a: Proficiency-, time-, and efficiency plots for participant B:



SDC1b: Proficiency-, time-, and efficiency plots for Participant F:



Supplementary Figure 1. Proficiency, time, and efficiency plots.