Learning Curves of Virtual Mastoidectomy in Distributed and Massed Practice

Steven Arild Wuyts Andersen, MD; Lars Konge, MD, PhD; Per Cayé-Thomasen, MD, DMSc; Mads Sølvsten Sørensen, MD, DMSc

**IMPORTANCE** Repeated and deliberate practice is crucial in surgical skills training, and virtual reality (VR) simulation can provide self-directed training of basic surgical skills to meet the individual needs of the trainee. Assessment of the learning curves of surgical procedures is pivotal in understanding skills acquisition and best-practice implementation and organization of training.

**OBJECTIVE** To explore the learning curves of VR simulation training of mastoidectomy and the effects of different practice sequences with the aim of proposing the optimal organization of training.

**DESIGN, SETTING, AND PARTICIPANTS** A prospective trial with a $2 \times 2$ design was conducted at an academic teaching hospital. Participants included 43 novice medical students. Of these, 21 students completed time-distributed practice from October 14 to November 29, 2013, and a separate group of 19 students completed massed practice on May 16, 17, or 18, 2014. Data analysis was performed from June 6, 2014, to March 3, 2015.

**INTERVENTIONS** Participants performed 12 repeated virtual mastoidectomies using a temporal bone surgical simulator in either a distributed (practice blocks spaced in time) or massed (all practice in 1 day) training program with randomization for simulator-integrated tutoring during the first 5 sessions.

**MAIN OUTCOMES AND MEASURES** Performance was assessed using a modified Welling Scale for final product analysis by 2 blinded senior otologists.

**RESULTS** Compared with the 19 students in the massed practice group, the 21 students in the distributed practice group were older (mean age, 25.1 years), more often male (15 [62%]), and had slightly higher mean gaming frequency (2.3 on a 1-5 Likert scale). Learning curves were established and distributed practice was found to be superior to massed practice, reported as mean end score (95% CI) of 15.7 (14.4-17.0) in distributed practice vs 13.0 (11.9-14.1) with massed practice ($P = .002$). Simulator-integrated tutoring accelerated the initial performance, with mean score for tutored sessions of 14.6 (13.9-15.2) vs 13.4 (12.8-14.0) for corresponding nontutored sessions ($P < .01$) but at the cost of a drop in performance once tutoring ceased. The performance drop was less with distributed practice, suggesting a protective effect when acquired skills were consolidated over time. The mean performance of the nontutored participants in the distributed practice group plateaued on a score of 16.0 (15.3-16.7) at approximately the ninth repetition, but the individual learning curves were highly variable.

**CONCLUSIONS AND RELEVANCE** Novices can acquire basic mastoidectomy competencies with self-directed VR simulation training. Training should be organized with distributed practice, and simulator-integrated tutoring can be useful to accelerate the initial learning curve. Practice should be deliberate and toward a standard set level of proficiency that remains to be defined rather than toward the mean learning curve plateau.
Virtual reality (VR) simulation training is increasingly being used in surgical training, including in temporal bone surgery where evidence of the efficacy and validity of VR simulation training of novices is emerging.\textsuperscript{1-6} Virtual reality simulation training can address some of the concerns and constraints of traditional dissection training and supervised surgery, for example, regarding patient safety issues, reduced working hours, and costs of facilities and instructors.

Acquisition of surgical skills requires repeated and deliberate practice regardless of the training modality,\textsuperscript{7,8} and the organization of this practice is also of importance: distribution of practice is more efficient for psychomotor skills learning compared with massed practice.\textsuperscript{9,10} Nevertheless, initial mastoidectomy training is frequently offered to trainees or residents as participation in a temporal bone course in places where facilities for frequent dissection are not readily available. Temporal bone skills courses are often organized with a single, massed block of practice. In contrast to this arrangement, VR simulation offers the possibility of repeated training spaced in time to the needs and at the convenience of the individual trainee.

The learning curve of repeated practice of any surgical procedure is crucial because of the implications for training and organization. For operating room performance, approximately 13 procedures are needed for technical competency in all the major steps of a complete mastoidectomy, with performances being assessed using a validated task-based checklist\textsuperscript{11}; for VR simulation training, performance (measured by a simulator-generated total score) plateaus after just 4 mastoidectomy procedures.\textsuperscript{12} Little is known regarding the learning curves in different practice programs and with other performance measurements, such as final product assessment.

In addition to the organization of surgical skills training, the learning environment is important and should be learner-centered and provide both tutoring\textsuperscript{6} and opportunity for self-directed learning.\textsuperscript{13} In line with this approach, some VR simulators can offer simulation-integrated tutoring and guidance. However, knowledge regarding the effect of this training approach on mastoidectomy performance and skills acquisition is limited\textsuperscript{4,14} and, to our knowledge, has not been investigated in association with the learning curve. It could be hypothesized that simulator-integrated tutoring accelerates skills acquisition and increases performance, but it is equally plausible that ongoing simulator-integrated tutoring impedes learning and that performance drops once tutoring is discontinued.

In this study, we wanted to explore the final product performance learning curves of mastoidectomy in VR simulation training with the aim of proposing an optimal self-directed program for initial mastoidectomy training. We therefore used a 2 × 2 study design to establish the effects of distributed and massed practice of mastoidectomy with and without initial simulator-integrated tutoring.

Methods

The ethics committee for the Capital Region in Denmark deemed this study exempt. All trainees provided written informed consent; participation was voluntary, and participants did not receive financial compensation.

VR Simulation Platform

The Visible Ear Simulator—a freeware VR temporal bone surgical simulator\textsuperscript{15,16} available for download from the Internet\textsuperscript{17}—was used in this study. A modified version 1.3, designed specifically for use in research, supported individual participant log-in with predefined conditions, such as automatic loading of the tutor function and autosaving of the virtual temporal bones. The simulator runs on a personal computer with a graphics card (GeForce GTX; Nvidia) and uses a haptic device (Geomagic Touch; 3D Systems) for drilling with force feedback. An optional simulator-integrated tutor function features greenlighting of the volume to be drilled in each step of a complete mastoidectomy in correspondence to an on-screen guide with text and illustrations.\textsuperscript{16}

Participants

Forty-three medical students from the Faculty of Health and Medical Science, University of Copenhagen, volunteered for participation in this study, which was organized as an extracurricular activity. We recruited students from any semester of study for participation, and the only exclusion criterion was previous VR simulation training of mastoidectomy. All participants were novices regarding temporal bone surgery because it is not part of the pregraduate curriculum. Included for study was a group of 21 (of 24) participants who completed time-distributed practice from October 14 to November 29, 2013, and a separate group of 19 (of 19) participants who completed massed practice on May 16, 17, or 18, 2014. Upon enrollment, participants completed a questionnaire on background demographics as well as computer and gaming experience.

Study Design

A 2 × 2 study design was used to investigate the learning curves of distributed and massed practice in VR simulation training of mastoidectomy, with randomization within these 2 practice groups for additional simulator-integrated tutoring during the initial 5 sessions (Figure 1). In the distributed practice...
group, blocks of 2 procedures were separated by at least 3 days; in massed practice, all of the procedures were completed during 1 course day.

In both practice programs, participants completed 12 repeated procedures consisting of a complete mastoidectomy with entry into the antrum and posterior tympanotomy. All participants had access to the on-screen step-by-step guide to the procedure but were otherwise self-directed. Before the first procedure, all participants received a 30-minute class lecture on the surgical anatomy of the temporal bone and a 5-minute hands-on navigation task in the simulator to familiarize the participants with the simulator controls as well as drilling using the haptic device.

Outcome and Statistical Analysis
The virtual temporal bone final products that were auto-saved at the end of the 30-minute sessions were assessed by 2 expert raters (P.C.-T. and M.S.S.) who were blinded to participant, session number, and practice and tutoring groups (examples of final product progression are shown in Figure 2). The assessment was done using a 26-item modified Welling Scale for final product analysis of performance previously detailed.18 The final product score was calculated as the mean of the scores assigned by the 2 raters. Data were analyzed using SPSS, version 22 (SPSS Inc) for MacOS X with analysis of variance, Pearson correlation coefficient, and nonlinear regression.

Results
The 21 participants in the distributed practice group were older (mean age, 25.1 years), more often male (15 [62%]), and had slightly higher mean gaming frequency (2.3 on a 1-5 Likert scale measured as never, yearly, monthly, weekly, and daily) compared with the 19 participants in the massed practice group (mean age, 23.6 years; male, 5 [26%]; mean gaming frequency, 1.6). However, no significant differences of final product performance scores were found in association with these and other background and computer or gaming factors. In distributed practice, blocks were spaced by a mean of 7.7 days, and there was no correlation of the number of days between sessions with the change in final product scores between sessions (Pearson r = 0.12; P = .23).

Learning curves for the mean final product performances were plotted for distributed and massed practice of VR simulation mastoidectomy with and without simulator-integrated tutoring in the first 5 sessions (Figure 3). During the
course of the 12 sessions of repeated practice, participants’ performance increased considerably in both practice groups (Table). As expected, distributed practice was found to increase performance significantly more than massed practice by the end of the training program.

The simulator-integrated tutor function also significantly improved participants’ performance ($P < .01$): the mean for all of the tutored sessions of both practice groups was 14.6 (95% CI, 13.9-15.2) compared with 13.4 (95% CI, 12.8-14.0) for the corresponding sessions of nontutored participants. In both practice groups, performance decreased when tutoring ceased; however, the performance dropped markedly more in the massed practice group than in the distributed practice group (Figure 3). With continued nontutored practice, the tutored participants achieved the same score by the final session as the nontutored participants of their respective practice group.

In nontutored massed practice, initial performance increased faster than in nontutored distributed practice, but the performance started to decline after the fourth session (Figure 3). In contrast to this finding, the performance of the nontutored participants in the distributed practice group asymptotically increased toward a plateau. Nonlinear regression was used to fit a sigmoid function for this learning curve: this analysis suggested a final product performance plateau of 16.0 (95% CI, 15.3-16.7) and that participants statistically reached this level by the ninth session.

### Discussion

In this study on the learning curves of mastoidectomy in repeated VR simulation training, we found that final product performance significantly increased in both practice programs but improved considerably more with distributed practice. In massed practice, the mean performance started to decline after the fourth session, whereas the mean performance in the distributed practice program gradually increased toward a plateau. We also found that simulator-integrated tutoring accelerated initial final product performance but at the cost of a drop in performance when tutoring was discontinued.

We used final product analysis of the mastoidectomy performance, which has limitations because it does not consider the process of the procedure—only the end result. However, final product analysis is a validated assessment tool of mastoidectomy performance and made the analysis of more than 450 performances feasible, which is a strength of this study. A limitation of this study was the difference in participant characteristics between the 2 practice groups: video gaming frequency and sex have been demonstrated to affect novice performance in VR surgical simulation. Randomization to practice groups was not possible for practical reasons but could have prevented these differences, even though the variations could not be demonstrated to be significantly associated with performance.

Consistent with established knowledge of other surgical procedures, distributed practice of mastoidectomy was found to be superior to massed practice. Learning and skills acquisition is dependent on memory consolidation, and spacing of practice allows this to occur. Distributed practice has been demonstrated to benefit complex psychomotor skills acquisition in VR simulation training of laparoscopy, including skills transfer. However, the optimal intertraining interval for skills practice and consolidation is still debated in the literature. An intertraining interval of 1 day was not sufficient for improving novice performance in non-VR simulation of myringotomy with ventilation tube insertion. In our study, the intertraining interval of distributed practice ranged from 3 to 16 days (mean, 7.7 days), which was sufficient for skills consolidation because the performance increased steadily toward a plateau. This consolidation could also be a mechanism explaining the protective effect of distributed practice on the drop in performance when simulator-integrated tutoring was discontinued: the performance drop in the distributed group was only half that observed in the massed practice group.

In a recent study, an automated and ongoing feedback system was used to improve the drilling performance of novices. However, ongoing simulator-integrated tutoring in repeated mastoidectomy practice has, to our knowledge, not been studied previously. We found that initial tutoring accelerated the learning curve and significantly increased final product performance compared with training using the on-screen instructional guide alone, but we also found that performance dropped considerably when the tutor function was discontinued. Nonetheless, performance at this point was still well above that of the first procedure, indicating that ongoing assistance by the tutor function did not hinder learning and could be useful for novices if applied correctly.

The main aim of this study was to propose an optimized program for self-directed VR simulation training of mastoidectomy on the basis of the different learning curves. The slope of the learning curve of the massed practice groups could suggest that 3 repeated procedures per block of training is more effective than just the 2 we allowed in the distributed practice program. In addition, simulator-integrated tutoring was useful and improved performance but should be used in con-
junction with distributed practice to minimize the performance drop after tutoring is discontinued. Together, a distributed training program with blocks of practice consisting of 1 simulator-tutored procedure followed by 2 untutored procedures could possibly be better but warrants investigation. Furthermore, the number of blocks necessary for adequate training is not evident at this time.

The learning curve of the nontutored participants in the distributed practice program most accurately reflects the true learning curve of VR mastoidectomy training. This learning curve was found to be a classic, negatively accelerated curve with progression (seemingly) toward a plateau after relatively few practices. This outcome is consistent with the findings of Nash et al, who demonstrated a plateau in automated simulator score performance after 4 performances in another VR simulator, with every mastoidectomy procedure separated by 1 week. Even though the overall score seemed to plateau at this point, time to completion of the task was found to decrease linearly without plateauing during the 6 repetitions investigated. In our study, time was fixed at 30 minutes to avoid time effects on performance, and the simulator could not be closed before the 30 minutes had passed. Nonetheless, based on observations during data collection, we believe that a substantial improvement in time to completion applies to our participants’ performances as well.

We found the mean learning curve of the distributed nontutored participants to plateau at a final product score of 16 of 26 possible points. Some technical limitations in the simulator could make it very hard to achieve the maximum final product score. However, this low-level performance at plateau cannot be attributed to such a ceiling effect or to the self-directed training program alone used in this study because the plateau is considerably lower than the highest observed performances: 4 performances (0.9%) were rated at 20 points or more, and 134 of all performances (29%) were rated at or above the mean plateau (two-thirds of these being nontutored performances).

The fitted curve suggested that the performances were within the level of the plateau at the ninth procedure, and any additional gain in the mean performance seemed only modest after the sixth repetition. The individual learning curves should be acknowledged at this point: different types of learning curves have been identified in surgical technical skills training with (1) a small group demonstrating immediate proficiency and therefore modest improvement, (2) the largest group improving at different rates with repeated practice, and (3) a small group underperforming without improvement during training. Even though our study interventions altered the learning curves, we found all of these patterns represented in our data, including in the nontutored distributed practice group. In VR laparoscopic simulation, some performances continued to increase beyond the initial plateau even after 30 sessions. We also found that the performance of some participants improved linearly without plateauing during the 12 procedures; this result substantiates that there is a high degree of variability in the length and amplitude of the learning curves of novices consistent with other reports. It is therefore possible that another order of magnitude of repetitions could continue to improve performance beyond the level of the initial plateau, with deliberate practice being key to avoid arrested development.

In general, the varying and highly individual learning curves undermine a predefined number of practice sessions, and training to the level of the mean plateau cannot ensure that the trainee has received the maximal benefit possible from VR simulation training. Such mastery learning instead calls for deliberate practice toward a predefined proficiency level. However, achieving this level is dependent on defining such a standard first and ideally using a validated performance assessment tool. Most reported instruments for mastoidectomy performance assessment, such as global rating scales and task-based checklists, rely on direct observation of performance, which, in contrast to final product analysis, can capture process and technique aspects. Nonetheless, all of these methods are dependent on the time-consuming rating of performance using trained experts, which seems less feasible in the context of deliberate practice and training toward a standard-set proficiency level. A future solution to this difficulty could be automated assessment based on simulator-gathered metrics with a standard setting of the level of proficiency.

Conclusions
Massed practice is often offered in temporal bone courses in places where training facilities do not provide an opportunity for repeated practice at the individual needs of the trainee. However, novices can acquire basic mastoidectomy competencies with self-directed VR simulation training, and repeated and distributed practice can support consolidation of these skills whereas massed practice is suboptimal. Simulator-integrated tutoring can accelerate the initial performance but should be applied in a way that facilitates optimal learning. Individual learning curves are highly variable; thus, the number of training sessions needed to ensure maximal benefit from VR simulation training greatly varies. However, deliberate practice should be aimed toward an evidence-based standard level of proficiency rather than the mean learning curve plateau. Such a predefined proficiency level has yet to be defined but could advantageously be established using future automatic and simulator-based assessment for real-time feedback and evaluation.
Obtained funding: Andersen, Sørensen. Administrative, technical, or material support: Konge, Sørensen.
Study supervision: Konge, Sørensen.
Conflict of Interest Disclosures: Dr Andersen has received an unrestricted grant from the Oticon Foundation for PhD studies. No other disclosures were reported.
Funding/Support: The development of the Visible Ear Simulator software has been financially supported by the Oticon Foundation.
Role of the Funder/Sponsor: The Oticon Foundation had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: Peter Trier Mikkelsen, DMSc (Alexandra Institute), developed the experimental version of the Visible Ear Simulator, and Sebastian Roed Rasmussen, BScMed, and Andreas Pagh Kohl, BScMed (Center for Clinical Education), assisted with data collection. There was no financial compensation.

REFERENCES