Accepted: 5 December 2016

DOI: 10.1111/eje.12254

ORIGINAL ARTICLE

WILEY

The effect of repeated testing vs repeated practice on skills learning in undergraduate dental education

S. Sennhenn-Kirchner[1,2] | Y. Goerlich[3] | B. Kirchner[1] | M. Notbohm[4] | S. Schiekirka[3] | A. Simmenroth[3,5] | T. Raupach[4,6]

[1]Department of Oral and Maxillofacial Surgery, University Medical Centre Goettingen, Göttingen, Germany

[2]Dental skills lab SINUZ, University Medical Centre Goettingen, Göttingen, Germany

[3]Student Training Centre of Clinical Practice and Simulation, University Medical Centre Goettingen, Göttingen, Germany

[4]Study Deanery, University Medical Center Goettingen, Göttingen, Germany

[5]Department of General Practice, University Medical Centre Goettingen, Göttingen, Germany

[6]Department of Cardiology & Pneumology, University Medical Centre Goettingen, Göttingen, Germany

Correspondence
Sabine Sennhenn-Kirchner, Department of Oral and Maxillofacial Surgery, University Medical Centre, Goettingen, Germany.
Email: se.ki@med.uni-goettingen.de

Abstract

Objectives: Recent studies in undergraduate medical education have demonstrated the advantage of repeated testing over repeated practice with regard to knowledge and skills retention. The aim of this study was to investigate whether this "testing effect" also applies to skills retention in undergraduate dental education.

Methods: In this prospective, randomised controlled trial, fourth-year dental students at Göttingen University Medical Centre participated in a training session on surgical suturing in winter term 2014/2015. Following this, they were either assigned to two sessions of additional skills training (group A) or two sessions of skills assessment with feedback (group B). These sessions were spaced over a period of 4 weeks. Skills retention was assessed in a summative objective structured clinical examination (OSCE) at the end of term, that is 6 months after the initial teaching session.

Results: A total of 32 students completed the study. With regard to suturing, OSCE performance was significantly better in group B than group A (81.9±13.1% vs 63.0±15.4%; P=0.001; Cohen's d=1.33). There was no significant OSCE performance difference in the two groups with regard to other learning objectives that were addressed in the end-of-term examination. Thus, the group difference was specific to suturing skills.

Conclusions: This is the first study to demonstrate that in dental education, repeated testing produces more favourable skills retention than repeated practice. Test-enhanced learning might be a viable concept for skills retention in undergraduate dentistry education.

KEYWORDS

dental education, objective structured clinical examination, OSCE, repeated assessment, retention of practical skills, testing effect

1 | INTRODUCTION

In undergraduate dental education, examinations are mainly used to assess whether students have met the learning objectives and include practical and oral examinations. Following training on simulation models in phantom heads, practical skills in clinical undergraduate dental courses are usually assessed by rating students' dental work on patients. However, the current standard is met by assessment of practical and communication skills via simulation in objective structured clinical examinations (OSCE).[1-8] Practical skills assessment in an OSCE via checklist and global rating provides reproducibility and reliability, both not being a given in the treatment of patients. "Summative" assessments are graded and usually create a strong incentive to prepare for a specific test, thus changing

student learning behaviour. "Formative" assessments, on the other hand, are not graded and cannot be failed. Instead, these examinations yield feedback for students, potentially guiding their future study behaviour. Recent research in medical education has revealed that formative testing does not only provide a means to check one's own knowledge. Moreover, it helps to enhance long-term storage of information according to the retrieval hypothesis underlying the direct testing effect, trying to retrieve information that has previously been stored in memory. This is confirmed by studies in various areas, including neurology[9-11] and internal medicine.[12] Most published studies focused on knowledge as the main learning outcome; however, some studies have demonstrated that the testing effect also applies to practical skills.[13,14]

In the field of dental education, Baghdady et al.[15,16] were able to show a main effect of testing condition compared to study condition on dental and dental hygiene students' comprehension of basic science mechanisms and diagnostic accuracy in dental radiology.

To the best of our knowledge, there have been no studies to date focusing on test-enhanced learning concerning practical skills in dentistry. To assess this, an appropriate test format is needed. One such format is the objective structured clinical examination that was introduced to dental education about 20 years ago.[3,5,6,8,17–19] More recently, a dental education OSCE focusing on suturing skills has been described.[20] The goal of our study was to assess the impact of repeated testing in comparison with repeated training on suturing skills retention in fourth-year dental students.

Our hypothesis was that repeated testing produces better suturing skills retention than repeated practice without testing.

2 | MATERIAL AND METHODS

At our institution, undergraduate dental education is divided into a pre-clinical and a clinical phase, both of which last two and a half years. Students start treating patients for the first time in a course in restorative dentistry during their third year. Simultaneously, they prepare for oral surgery courses in simulation sessions without practical end-of-term tests, for example suturing courses on pig's jaws. During their fourth and fifth year, they participate in oral surgical courses involving patients.

2.1 | OSCE design and piloting

In summer 2014, a new form of assessment was designed for the first time in the dental clinic in Göttingen. It consisted of five oral surgical stations (informed consent, postoperative consultation following a tooth extraction, local anaesthesia, tooth extraction, suturing). The suturing station demanded one single traumatic suture, threading included, as well as one non-traumatic suture. The OSCE was piloted in October 2014 to evaluate test quality criteria. The reliability of the checklists used ranged from 0.48 to 0.81 (see Table 1). All raters attended standard rater training sessions to ensure standard valuation and calibration.

2.2 | The study

In winter term 2014/2015, all 36 fourth-year dental students enrolled in the oral surgery course were invited to participate in a prospective randomised controlled trial. Students providing written consent were stratified according to sex and medical qualification as students who have previously completed vocational training as dental assistants or nurses may already have more experience than dental students without additional qualifications. Students were then randomised into one of two groups (see Figure 1). Following a practical training session providing an intensive 4-hour course with focus on suturing skills in summer 2014, all students participated in two booster sessions, 5 and 6 months after initial training, respectively. During these sessions, students in the practice group (A) were able to practice suturing skills, sharing one suture simulation pad between two students. All students were supervised by two oral surgeons who answered questions and corrected mistakes but did not volunteer feedback to all students. Supervisors were specifically trained to ensure that all sessions were standardised with regard to length, content and teacher behaviour. Students in the testing group (B) experienced a different setting. Here, every student entered a room on her/his own and took a formative assessment. A supervisor watched the student perform the suturing task and provided a rating on a sheet that was given to the student. No specific feedback was given. Supervisors were the same two oral surgeons in both groups, trained exactly the same way as the blinded OSCE raters. Topic, equipment and time available were comparable in groups A and B. Every student in group A spent 10 minutes for training. Participants in group B were offered eight minutes testing time and two minutes to receive a short written feedback. Retention of suturing skills as well as four other skills that had been taught throughout summer term 2014 was assessed in an objective structured clinical examination (OSCE) in February 2015. The study results were gained in this end-of-term OSCE. The suture station evaluated student's performance in threading, handling of the acutenaculum, hand and instrumental knotting, overall hygienic performance, uniformity and risk of injuries' avoidance.

2.3 | End-of-term assessment

The results of the study were gained in the end-of-term OSCE in February 2015. The suture station for instance evaluated students' performance in threading, handling of the acutenaculum, hand and instrumental knotting, overall hygienic performance, uniformity and risk of injuries' avoidance.

2.4 | Data collection and statistical analysis

Data were recorded electronically referring to the study's end point ("performance in the end-term OSCE") using tablets (Apple iPad Air; Apple Inc., Cupertino, CA, USA). The server application "tOSCE" (Umbrella Consortium for Assessment Networks; UCAN) was used to store data and export them to a CSV file.

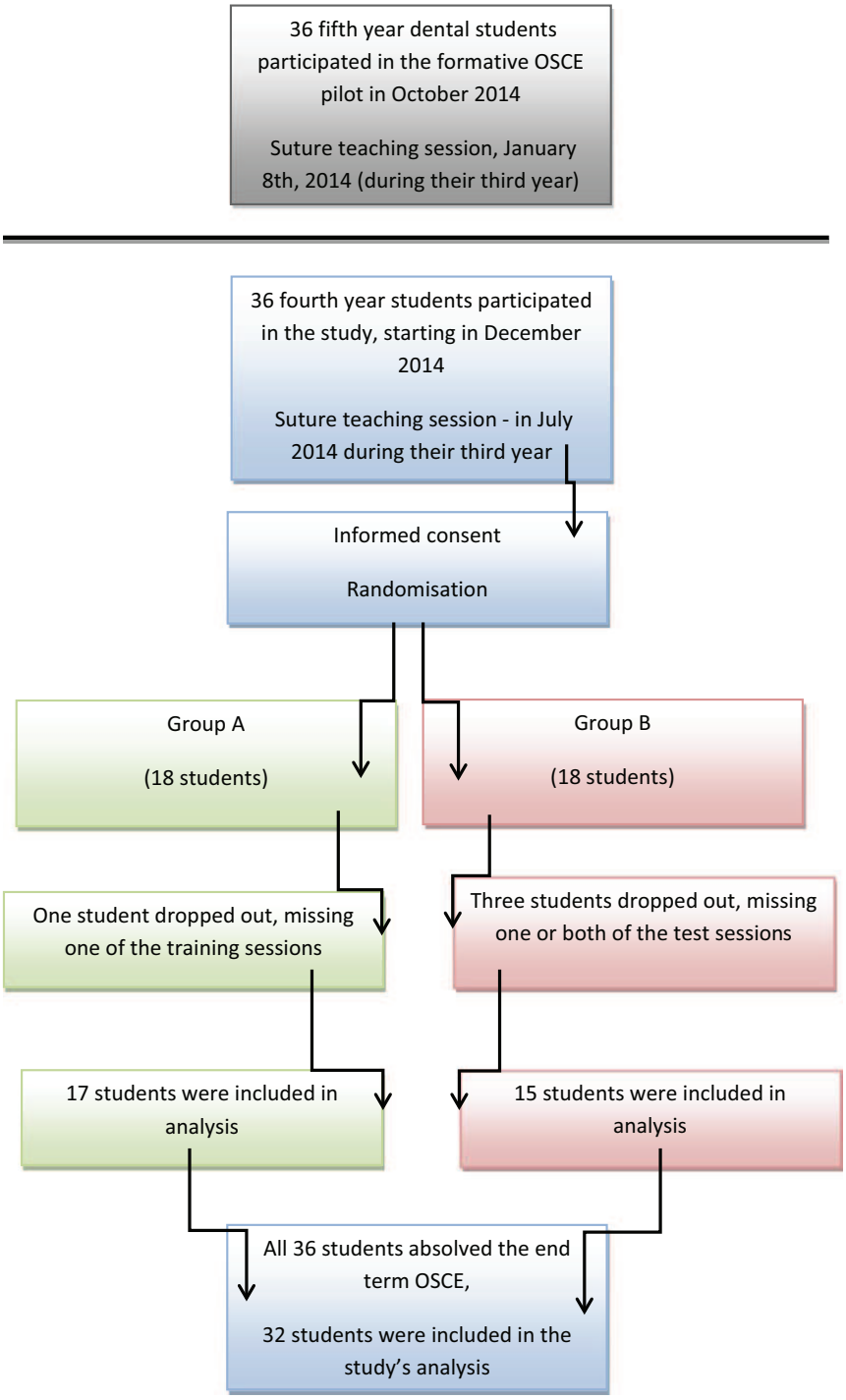

FIGURE 1 The OSCE pilot was performed in October 2014 including voluntary fifth-year dental students. These students attended an one-day suture course in January 2014. A total of 36 fourth-year students were invited to participate in the trial from November 2014 up to end of term in February 2015, and 32 students completed all parts of the trial. This cohort participated in an one-day suture course in July 2014 [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Descriptive analysis (mean ± standard deviation) and internal consistency of OSCE stations (pilot phase October 2014)

Station	N	Per cent score	Cronbach's α
Informed consent	36	64.6±20.8	0.725
Postoperative consultation	36	71.5±19.0	0.811
Local anaesthesia	36	63.4±16.7	0.481
Tooth extraction	36	64.2±14.2	0.530
Suture	36	67.9±13.6	0.621

Results for all OSCE stations were made up of two components: the first component contained a number of specific outcomes that were marked as 0 ("incorrect"), 0.5 ("partially correct") and 1 ("correct"). For each student, a sum score for this component was calculated, and this contributed by 70% to the total result. The second component was one single overall rating on a scale anchored by 1 ("excellent")

TABLE 2 OSCE results per station (mean per cent score±standard deviation) for the total sample as well as for the two study groups separately. *P* values refer to independent *t* tests

Station	All students (n=32)	Cronbach's α	Group A (n=17)	Group B (n=15)	*P* value
Informed consent	61.0±15.6	0.655	58.1±15.5	64.4±15.5	0.262
Postoperative consultation	70.4±18.0	0.734	73.1±15.0	67.4±20.9	0.381
Local anaesthesia	78.2±14.6	0.494	75.7±17.9	81.7±9.3	0.245
Tooth extraction	79.3±11.3	0.462	76.7±12.5	82.5±8.9	0.149
Suture	71.9±17.8	0.745	63.0±15.4	81.9±13.1	0.001
Total	72.2±8.9	0.752	69.2±8.8	75.5±8.1	0.046

and 5 ("very poor"). These marks were recoded as follows: 1=1 point, 2=0.75, 3=0.5, 4=0.25 and 5=0 points. The overall rating contributed by 30% to the total result. Total score is reported as a percentage. Internal consistency of each checklist was assessed by calculating Cronbach's α. Results are presented as mean±standard deviation. Differences between the two groups were assessed using independent *t* tests. Statistical testing was performed with IBM SPSS Statistics version 21.

2.5 | Ethics approval

The study protocol was approved by the institutional review board of Göttingen University Medical Centre (application number 5/11/14).

3 | RESULTS

A total of 36 students volunteered to participate in the trial, but four had to be excluded due to incomplete attendance at intervention or control sessions. Table 2 illustrates the results of all OSCE stations in per cent, differentiated into group A and group B. The results of group B seem to exceed group A, but a significant difference between the groups was only observed for the suture station (81.9±13.1% vs 63.0±15.4%; *P*=0.001; Cohen's *d*=1.33).

4 | DISCUSSION

The idea for our study was based on the results of an oral surgery OSCE recently introduced in Göttingen dental clinic. We were able to perform the piloting of five stations and corresponding checklists with the support of 36 voluntary fifth-year students. By means of the results of the pilot study, OSCE stations and checklists were revised and adapted to fit test quality criteria.

The participants of the main study were all 36 fourth-year dental students enrolled in the regular oral surgery course including an end-of-term OSCE. Four students had to be excluded from the study as having missed one or both of the study interventions due to personal reasons or caused by illness. Study participation was voluntarily, and OSCE results of the suture station were declared to be formative.

The acquisition of clinical-technical skills is central to undergraduate dental education. Studies have shown that summative assessments are potent drivers of student learning activity.[2,4,21] More recently, test-enhanced learning facilitated by formative assessments has received increasing attention as a means of enhancing knowledge and skills retention, and this study is one of the first to demonstrate the testing effect in dental education.[4,5,9,10,13–16,22] A specific finding of the current study is that test-enhanced learning also works when initial teaching and booster testing sessions are spaced considerably: in contrast to earlier trials using shorter delays, the first testing session occurred more than 5 months after initial teaching.[2,10,15,16,23] We still found an effect of testing on examination performance 1 month after the final booster session, indicating that spacing between initial learning and testing may be less important than (short) spacing between booster sessions and eventual skills assessment.

On the one hand, better retention should be gained by repetition, which means more suturing in the following surgical course. On the other hand, there is a shortage of patients suitable for student practice as well as of teaching time, so that this form of teaching unfortunately could not be offered regularly. Approaches from other areas of medicine teaching should therefore be explored in terms of their transferability to oral surgical studies.

Roediger and Butler stated retrieval practice, as occurs during testing, often exceeds studying in its effects on long-term retention of knowledge.[24] This statement is supported by Dobson and Linderholm, who found testing to be superior to independent study.[23] Kromann et al.[13,14] state that in addition to the extrinsic effects of assessment and examinations on students' study habits, testing can have an intrinsic effect on the recall of studied material. Their prospective, controlled and randomised intervention study was conducted to determine whether testing as the final activity in a resuscitation skills course (intervention) increases learning compared with an equal amount of time spent practicing the skill (control). Their results are especially interesting in view of the need to maximise learning outcomes from costly simulation-based courses. They evaluated significantly higher learning outcomes in the intervention group (n=41; mean score 82.8%, 95% confidence interval [CI] 79.4-86.2) compared with the control group (n=40; mean score 73.3%, 95% CI: 70.5-76.1) with *P*<0.001, effect size 0.93.[13] In our study, we were able to show a comparable effect in a suture setting (Cohen's *d*=1.33). In 2009, Wood assessed in course testing as being beneficial for students to stimulate the training process with high potential for long-term retention. He found that assessment not only drives, but even helps learning,

and the results of our study show the same for practical skills retention.[25] A study by Carpenter in the same year indicates that activation of elaborative retrieval (whereby testing shows better results than restudying) may be one of the mechanisms underlying the testing effect.[26]

4.1 | Limitations

Compared with undergraduate medical education, the number of dental students is distinctly smaller. Therefore, the main limitation of our study is a small sample size, according to the small number of students attending the same dental class at our institution. For this reason, we were unable to include an equivalent control group.

However, the sample was large enough to detect the large effect elicited by repeated testing compared to repeated practice (Cohen's d=1.33). Given that the student cohort was small and students interact with each other a lot, contamination between the two groups (i.e students in one group informing students in the other group about the content of practice/testing sessions) is a potential threat to the validity of our results. We found no evidence of students in the control group practicing more to make up for the fact they did not receive the intervention, but even if such compensation had taken place, it did not override the observed effect.

Comparing both groups on use of resources, testing group requires more teaching time due to one-on-one supervision.

5 | CONCLUSION

This is the first study to demonstrate that in dental education, repeated testing produces more favourable skills retention than repeated practice. The results confirmed our hypothesis that students who had experienced repeated testing would score more points and achieve higher grades in an end-of-term OSCE than students who repeatedly practiced the same skill. One particular feature of this study was the long delay between initial teaching and the first booster (testing) session. Test-enhanced learning might be a viable concept for skills retention in undergraduate dentistry education.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

REFERENCES

1. Eberhard L, Hassel A, Bäumer A, et al. Analysis of quality and feasibility of an objective structured clinical examination (OSCE) in preclinical dental education. *Eur J Dent Educ*. 2011;15:172–178.
2. Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356:387–396.
3. Govindan VK. Enhancing communication skills using an OSCE and peer review. *Med Educ*. 2008;42:535–536.
4. Kiehl C, Simmenroth-Nayda A, Goerlich Y, et al. Standardized and quality-assured video-recorded examination in undergraduate education: informed consent prior to surgery. *J Surg Res*. 2014;191:64–73.
5. Landes CA, Hoefer S, Schuebel F, et al. Long-term prospective teaching activity of practical skills training and a first OSCE in cranio maxillofacial surgery for dental students. *J Craniomaxillofac Surg*. 2014;42:e97–e104.
6. Manogue M, Brown G. Developing and implementing an OSCE in dentistry. *Eur J Dent Educ*. 1998;2:51–57.
7. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65:63–67.
8. Schoonheim-Klein ME, Habets LLMH, Aartmann IHA, van der Vleuten CP, Hoogstraaten J, van der Velden U. Implementing an Objective Structured Clinical Examination (OSCE) in dental education: effects on students' learning strategies. *Eur J Dent Educ*. 2006;10:226–235.
9. Larsen DP, Butler AC, Roediger HL 3rd. Test-enhanced learning in medical education. *Med Educ*. 2008;42:959–966.
10. Larsen DP, Butler AC, Roediger HL 3rd. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ*. 2009;43:1174–1181.
11. Larsen DP. Picking the right dose. The challenges of applying spaced testing to education. *J Grad Med Educ*. 2014;6:349–350. doi: 10.4300/JGME-D-1400170.1.
12. Schmidmaier R, Ebersbach R, Schiller M, Hege I, Holzer M, Fischer MR. Using electronic flashcards to promote learning in medical students: retesting versus restudying. *Med Educ*. 2011;45:1101–1110.
13. Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Med Educ*. 2009;43:21–27.
14. Kromann CB, Bohnstedt C, Jensen ML, Ringsted C. The testing effect on skills learning might last 6 months. *Adv Health Sci Educ Theory Pract*. 2010;15:395–401.
15. Baghdady MT, Carnahan H, Lam EW, Woods NN. Test enhanced learning and its effect on comprehension and diagnostic accuracy. *Med Educ*. 2014;48:181–188.
16. Baghdady MT, Carnahan H, Lam EW, Woods NN. Dental and dental hygiene students' diagnostic accuracy in oral radiology: effect of diagnostic strategy and instructional method. *J Dent Educ*. 2014;78:1279–1285.
17. Höfer SH, Schuebel F, Sader R, Landes C. Development and implementation of an objective structured clinical examination (=SCE) in CMF-surgery for dental students. *J Craniomaxillofac Surg*. 2013;41:412–416.
18. Karpicke JD, Blunt JR. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*. 2011;331:772–775.
19. Yudkowsky R, Alseidi A, Cintron J. Beyond fulfilling the core competencies: an objective structured clinical examination to assess communication and interpersonal skills in an surgical residency. *Curr Surg*. 2004;61:499–503.
20. Macluskey M, Durham J, Balmer C, et al. Dental student suturing skills: a multicenter trial of a checklist-based assessment. *Eur J Dent Educ*. 2011;15:244–249.
21. Buss B, Kratter M, Möltner A, Weyrich P, Juenger J, Nickendei C. Can the ‚assessment drives learnin' effectbe detected in clinical skills training? –Imlications for curriculum design and resource planning. *GMS Z Med Ausbild*. 2012;29:DOC70.
22. Dobson JL, Linderholm T. The effect of selected "desirable difficulties" on the ability to recall anatomy information. *Anat Sci Educ*. 2015;8:395–403.
23. Raupach T, Brown J, Anders S, Hasenfuss G, Hanrendza S. Summative assessments are more powerful drivers to students learning than resource intensive teaching formats. *BMC Med*. 2013;11:61.

24. Roediger HL 3rd, Butler AC. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci*. 2011;15:20–27.
25. Wood T. Assessment not only drives learning, it may also help learning. Comment in: Do we need to train assessors? *Med Educ*. 2009;43:5–6.
26. Carpenter SK. Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J Exp Psychol Learn Mem Cogn*. 2009;35:1563–1569.

How to cite this article: Sennhenn-Kirchner S, Goerlich Y, Kirchner B, et al. The effect of repeated testing vs repeated practice on skills learning in undergraduate dental education. *Eur J Dent Educ*. 2017;00:1–6. doi:10.1111/eje.12254.