

## Research Article

# Improving Surgical Anatomy Teaching

Doomernik DE<sup>1\*</sup>, van Linge A<sup>1</sup>, Donders R<sup>2</sup>, Vorstenbosch MATM<sup>1</sup>, van Goor H<sup>3</sup> and Ruiter DJ<sup>1</sup>

<sup>1</sup>Department of Anatomy, Radboud University Medical Center, The Netherlands

<sup>2</sup>Department for Health Evidence, Radboud University Medical Center, The Netherlands

<sup>3</sup>Department of Surgery, Radboud University Medical Center, The Netherlands

\*Corresponding author: Doomernik DE, Department of Anatomy, Radboud University Medical Center, Postbox 9101 6500 HB Nijmegen, The Netherlands

Received: October 05, 2015; Accepted: December 04, 2015; Published: December 07, 2015

## Abstract

**Background:** Medical students in clinical settings may have difficulty to recall and apply anatomical knowledge in diagnostic reasoning and problem-solving. In this double blind prospective randomized trial we aimed to assess A) the learning effect of anatomy teaching in the Surgery Clerkship Preparation Course (SCPC) and B) whether offering an anatomical pretest has an additional learning effect.

**Methods:** In a 10 months period, 10 SCPC groups were randomly assigned in 2 study arms. The intervention arm (5 groups, n=128) received an anatomical pretest at day 1. The control arm (5 groups, n=135) received a sham pretest. All students participated in anatomy classes at day 2-4 and underwent an anatomical posttest at day 15. Pre- and posttest scores were corrected for item difficulty (Modified-Angoff method). Pre- and posttest scores of the intervention group were compared to assess the learning effect of anatomy teaching in the SCPC (Random Effects Model) and posttest scores of both groups were compared to assess the effect of pretesting.

**Results:** A significant improvement on posttest performance was seen in the intervention arm ( $p < 0.0005$ ). No significant differences in posttest scores were seen ( $p = 0.857$ ) between the intervention and control group.

**Discussion:** This study demonstrates a significant learning effect of anatomy teaching in SCPC. However, no adjuvant effect of pretesting on directing students to subsequent learning was seen. These findings underscore the positive value of anatomy teaching in a SCPC and suggest that the pretest offered does not add to the learning effect in the current study design.

**Keywords:** Anatomy; Medical education; Subsequent learning; Pretesting

## Abbreviations

SCPC: Surgery Clerkship Preparation Course; Radboudumc: Radboud University Medical Center; MCQ's: Multiple-Choice Questions; TM: Trimmed Averages; CF: Correction Factor

## Introduction

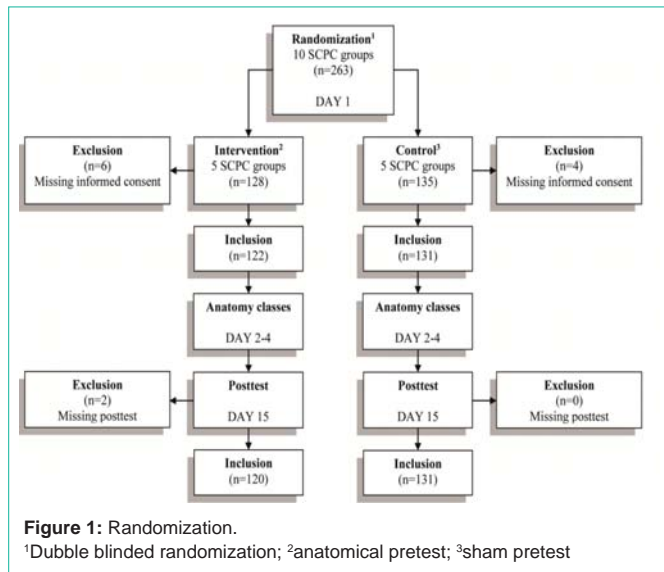
In most educational settings, tests have traditionally been used as assessment tools to evaluate students learning. Testing has been shown to increase long-term retention, i.e. 'test-enhanced learning'. Test-enhanced learning refers to the fact that taking an initial test on previously studied material produces better retention over time than when not tested on that material. Frequent testing results indirectly in better retention due to increased efficiency of subsequent study time and improvement of study strategies. This phenomenon is called the testing-effect. Research suggests that the testing effect is driven by retrieval practice and not by repeated study [1,2].

Testing before study (i.e. pretesting) has shown to improve subsequent learning of pretested and non-pretested information and may lead to better recall [3,4]. The pre-testing effect is based on the finding that students score better in a pretest condition than in an extended learning condition on posttest performance. Pre-testing may be beneficial because it encourages a more active participation in learning, directs attention to important information and makes it easier to distinguish what is important to learn and what not even

when test results are below average. However, a poor result on a pretest may also negatively affect the learning process [4].

The Radboud University Medical Center (Radboudumc) provides a problem-oriented, horizontally and vertically integrated medical curriculum. In the first 3 years of the undergraduate curriculum (Bachelor) the courses are organized by interdisciplinary themes, i.e. basic science and subsequent organ specific themes. During the last 3 years, the clinical phase (Master of Science), students follow a tailor made preparatory course each time they enter a clerkship of the clinical rotation in one of the major clinical disciplines, according to "the just in time" learning principle [5,6]. This "just in time" learning principle is based on the knowledge that spaced and vertical education, integration of basic sciences in an undergraduate and continuation of basic sciences in the later years of a medical curriculum, using problem based learning principles, improves knowledge retention and clinical reasoning [5,6].

We have observed that medical students in preparatory courses and in clinical settings have difficulty to recall, understand and apply their anatomical knowledge in diagnostic reasoning and problem-solving, especially regarding the surgery clerkship. We assume that anatomy teaching in a 3 weekly Surgery Clerkship Preparation Course (SCPC) prepare students' well for their surgery clerkship. However, the learning effect of anatomy teaching in the SCPC has never been studied before. In the present study we aimed to assess in a



double blinded prospective randomized trial A) the learning effect of teaching anatomy in the SCPC and B) whether offering an anatomical pretest has an additional learning effect.

## Methods

### Design and setting

During a period of 10 months, 10 SCPC groups (n=263) were randomly assigned into two arms of equal numbers. Students in the intervention arm (5 groups, n=128) were given an anatomical pretest at day 1, anatomy classes at day 2-4 and an anatomical posttest at day 15. Students in the control arm (5 groups, n=135) were given a sham pretest at day 1, anatomy classes at day 2-4 and a posttest at day 15 (Figure 1). Anatomy teaching was provided each month by the same two independent tutors. Both control and intervention group attended the same anatomy classes on applied anatomy of thorax, abdomen, upper and lower extremities. Basic anatomical knowledge is retrieved and students are trained in applying knowledge in diagnostic reasoning and problem solving through lectures, task-driven non-directed self-study, tutorials and guided station-based cadaver practicals. The pretest consisted of 10 Multiple-Choice Questions (MCQ's) with a maximum of four alternative answers at the level of the Bachelor Medical curriculum. The posttest consisted of 10 MCQ's with a maximum of four alternative answers at the knowledge level that is expected after completing the anatomy classes in SCPC. The sham pretest included 11 statements about surgeons and the upcoming surgical clerkship on a five-point Likert scale (agree versus disagree). The multiple-choice questions and statements were formulated by one of the authors (AvL), and validated on content by an independent expert panel consisting of an anatomist, surgeon and three medical educationalists. The students were given 15 minutes to finish the assessment and encouraged to answer each question. An independent observer prevented peeking. No feedback was given. Both tutors and researcher were blinded for the randomization arm and the content of the questions of the pre- and posttest.

### Data and statistical analysis

Test results were scored blindly. Each correct answer yielded 1 point with a maximum score of 10 points. No correction for

**Table 1: Correction for test difficulty with the Modified-Ang off method.**

	Pretest	Posttest
Formula	$TM_{pretest} / (TM_{pretest} + TM_{posttest})$	$TM_{posttest} / (TM_{pretest} + TM_{posttest})$
TM <sup>1</sup> (range 1-5)	2.99	3.68
CF <sup>2</sup> (range 0-1)	0.448	0.552

<sup>1</sup>TM: Trimmed Average; <sup>2</sup>CF: Correction Factor for item difficulty

guessing was used. Data were noted and analyzed using SPSS® for Windows® version 18.0 (SPSS, Chicago, Illinois, USA). The difficulty of the MCQ's was validated by using the Modified-Angoff method (Table 1). An expert panel consisting of 6 independent anatomists and educationalists reviewed the difficulty of the pre- and posttest MCQ's by predicting the performance of a minimally competent student on each item of both tests using a six-point Likert scale (very easy to extremely difficult). Overall test difficulty was calculated as sum of item scores and expressed as Trimmed Averages (TM). The Correction Factor (CF) for test difficulty was used to compare pre- and posttest scores (Table 1). A Random Effects Model was used to test the learning effect in the intervention group and to test the effect of pretesting. A random study group dependent intercept was added to correct for dependencies caused by the fact that randomization took place on study group level and that education took place in study groups. We assumed an unstructured covariance matrix between the residuals to account for repeated measurements of participants. The Pearson correlation coefficient was used to determine the relation between pre- and posttest scores. For all tests, p<0.05 was considered statistically significant (Table 1).

### Ethical considerations

At the start of this study, there was no formal ethical approval process for medical education research in the Netherlands. Information concerning possible risks for the students, the equitability of the selection, the guarantee of privacy and confidentiality and the procedure on informed consent was provided oral and written. Formal approval was obtained from the course coordinator. Participation was on a voluntary basis, participants could withdraw at any time and the test results did not affect the academic progress of the participants in any way. Identifying information were linked to a study number and made anonymous. Results were not disclosed to the course coordinator. The ethical principles of the World Medical Association Declaration of Helsinki (2008) were taken into account.

## Results

A total of 251 students were included for analysis, 128 students in the intervention and 135 in the control group (Figure 1). The Modified-Angoff method demonstrated trimmed averages for test difficulty of 2.99 for the pretest and 3.68 for the posttest which corresponds to a correction factor for difficulty of 0.448 and 0.552, respectively (Table 1).

### Learning effect

In the intervention group of 128 students, the mean pretest score was 4.8 (SD±1.8) and the mean posttest score was 5.0 (SD±1.9). After correction for test difficulty, pre- and posttest scores were 2.17±0.11 and 2.78±0.12 respectively. A significant improvement on posttest performance was seen after correction for test difficulty in the intervention group (p= < 0.0005) (Table 2).

**Table 2:** Corrected and uncorrected mean and standard deviation of pre- and posttest scores in the intervention and control arm.

	Pretest scores (mean $\pm$ SD)		Posttest scores (mean $\pm$ SD)		P-value (95% CI)
	Intervention	4.8 $\pm$ 1.8 <sup>1</sup>	2.17 $\pm$ 0.11 <sup>2</sup>	5.0 $\pm$ 1.9 <sup>1</sup>	
Control	Sham	Sham	5.1 $\pm$ 1.5 <sup>1</sup>	2.82 $\pm$ 0.18 <sup>2</sup>	
P-value (95% CI)				0.857 <sup>4</sup>	

SD: Standard Deviation; CI: Confidence Interval

<sup>1</sup>Mean score (range 1-10); <sup>2</sup>Mean score corrected for item difficulty; <sup>3</sup>P-value for comparing pre- and posttest scores in the intervention arm, which corresponds to the learning effect of anatomy teaching in the SCPC; <sup>4</sup>P-value for comparing posttest scores of the intervention and control arm, which corresponds to the pretesting effect.

### Pretesting effect

The mean posttest score of students in the intervention arm was 5.0 (SD $\pm$ 1.9) compared to 5.1 (SD $\pm$ 1.5) in the control arm. Posttest scores did not significantly differ between the intervention and control group ( $p=0.857$ ). No correlation was seen between pre- and posttest scores ( $r=-0.003$ ) (Table 2).

## Discussion

We were interested in the learning effect of anatomy teaching in the SCPC. As we expected, a significant learning effect was seen in the intervention group when comparing pre- and posttest scores. Another goal of our study was to evaluate pretesting as a tool to direct students to subsequent learning. Contrary to our expectations and to the literature, this study did not demonstrate an adjuvant learning effect of pretesting [3,4].

The significant learning effect might be explained by the continuous rehearsal of anatomical knowledge in our problem-oriented spiralized curriculum [7,8]. Previous studies have demonstrated that spaced education, vertical integration and problem-based learning principles improve knowledge retention and clinical reasoning [5,6]. Further, it has been shown that (pre)testing improves knowledge retention [2-4]. The efficacy of a test in promoting retention is influenced by the format in which the test is given [1] and by the content of the questions [9]. MCQ tests are objective, easy, quick, reliable in the discrimination between high and low performance students, enhance retention of the material tested and are therefore most commonly used as assessment method in medical education [9]. However, MCQ tests are limited by the cueing effect [9], i.e. less active recall of knowledge and retention of incorrect information in the form of lure items (wrong answers causing false memories) [10]. The value of using a MCQ test as a learning tool depends on the number of lure items in the test and the amount of study prior to the test [10]. Students in the intervention group were exposed to more lure items than the control group; however no negative effect of MCQ testing was seen on posttest performance. Further, the total exposure time to the subject matter and posttest performance were equal for both groups, suggesting that the anatomy classes effected learning more than the pretest. This finding is further supported by the significant learning effect seen in the intervention group between pre- and posttest scores.

The efficacy of a test in promoting retention depends also on the reliability and validity of the test [9]. The wide range of topics in the pre- and posttest (i.e. thorax, abdomen and the upper and lower extremities) may have contributed to the low inter-item

correlation. Since the number of questions is directly related to the reliability of a test, we assume that more extensive pretesting with multiple questions directed at each learning objective would more likely facilitate students in learning. The use of images in a response format also have effect on item difficulty and item discrimination and thus have implications on the validity of the test [11]. Both pre- and posttest consisted of context-free questions with and without images. Context-free questions reflect factual knowledge instead of diagnostic reasoning and problem-solving [9]. As we noted that our students have problems in applying anatomical knowledge in diagnostic reasoning, in retrospect we should have combined the context-free questions with context rich questions or case scenario's.

This study is strengthened by the double blinded prospective randomized controlled design and the use of the random effect model to correct for group effects making selection/sampling bias, information bias, confounding bias unlikely. Tutors were not involved in the research and the investigators were not involved in the SCPC course. It was thus unlikely that students intended to please the investigators or the tutors or vice-versa, minimizing response bias. Further, it was unlikely that students recalled their anatomical knowledge from different sources, because students were confronted with the study at the first day of SCPC. The key of the SCPC course is to acquire knowledge by self-directed learning. Students are assessed by level of participation and not by an examination, that's why the posttest results reflect actual knowledge. The lack of intrinsic motivation due to the formative test setting and the anatomical peak load on day 2-4, may have induced loss of focus by the students on the anatomy part and are possible explanations for the low correlation between pre- and posttest scores. Since feedback is known to improve the effectiveness of testing in generating long term retention [2], the lack of feedback on the pretest might have negatively influenced the pretesting effect. Further research should focus on offering frequent interim assessments during the SCPC and the surgery clerkship to enhance spaced anatomy education. This may further improve the educational contribution of anatomy teaching in the SCPC that was already shown to have a significant learning effect.

## Conclusion

Anatomy teaching as part of a surgical clerkship preparatory course appears to have a significant learning effect. However, pretesting of anatomical knowledge did not direct students to subsequent learning. These findings underscore the positive value of anatomy teaching in a SCPC and suggest that the pretest offered does not add to the learning effect in the current study design.

## References

- Larsen DP, Butler AC, Roediger HL. Test-enhanced learning in medical education. *Med Educ.* 2008; 42: 959-966.
- Larsen DP, Butler AC, Roediger HL. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ.* 2009; 43: 1174-1181.
- Kornell N, Hays MJ, Bjork RA. Unsuccessful retrieval attempts enhance subsequent learning. *J Exp Psychol Learn Mem Cogn.* 2009; 35: 989-998.
- Richland LE, Kornell N, Kao LS. The pretesting effect: do unsuccessful retrieval attempts enhance learning? *J Exp Psychol Appl.* 2009; 15: 243-257.
- Kerfoot BP, Fu Y, Baker H, Connelly D, Ritchey ML, Genega EM. Online spaced education generates transfer and improves long-term retention of

- diagnostic skills: a randomized controlled trial. *J Am Coll Surg.* 2010; 211: 331-337.
6. Norman GR, Schmidt HG. The psychological basis of problem-based learning: a review of the evidence. *Acad Med.* 1992; 67: 557-565.
  7. Harden RM, Davis MH, Crosby JR. The new Dundee medical curriculum: a whole that is greater than the sum of the parts. *Med Educ.* 1997; 31: 264-271.
  8. Bergman EM, van der Vleuten CP, Scherpbier AJ. Why don't they know enough about anatomy? A narrative review. *Med Teach.* 2011; 33: 403-409.
  9. Schuwirth LW, van der Vleuten CP. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ.* 2004; 38: 974-979.
  10. Roediger HL, Marsh EJ. The positive and negative consequences of multiple-choice testing. *J Exp Psychol Learn Mem Cogn.* 2005; 31: 1155-1159.
  11. Vorstenbosch MA, Klaassen TP, Kooloos JG, Bolhuis SM, Laan RF. Do images influence assessment in anatomy? Exploring the effect of images on item difficulty and item discrimination. *Anat Sci Educ.* 2013; 6: 29-41.