

# The reliability and predictive validity of a sixth-semester OSPE in conservative dentistry regarding performance on the state examination

## Abstract

**Introduction:** The aim of this study was to ascertain whether the testing format of an OSPE (Objective Structured Practical Examination) in conservative dentistry (sixth semester) predicts the scores on the practical section of the state examination (11<sup>th</sup> semester) in the same subject. Taking general student profiles into consideration (score on the school-leaving exam [Abitur], score on the preliminary exam in dental medicine [Physikum], length of university study, cohorts, and sex), we also investigated if any correlations or differences exist in regard to the total and partial scores on the OSPE and the corresponding state examination.

**Methods:** Within the scope of this longitudinal retrospective study, exam-specific data spanning 11 semesters for dental students (N=223) in Frankfurt am Main were collected and analyzed. Statistical analysis was carried out by calculating Spearman rank correlations, partial correlations, Pearson's correlation coefficients, and multiple regressions (SPSS Statistics 21, IBM Corporation, New York).

**Results:** The results show that the OSPE (Cronbach's  $\alpha=.87$ ) correlates with level of success on the practical section of the state exam in conservative dentistry ( $p=.01$ ,  $r=.17$ ). Length of university study also emerged to correlate significantly with the state exam score ( $p=.001$ ,  $r=.23$ ). Together, these two variables contribute significantly to predicting the state exam score ( $p=.001$ ,  $R^2=.076$ ). This was seen extensively among female students. It was also discovered that these female students had higher school-leaving exam scores than male students ( $F=6.09$ ,  $p=.01$ ,  $\eta^2=.027$ ), and that a significant correlation between scores on the Physikum (preliminary exam in dental medicine) and OSPE scores existed only for male students ( $r=.17$ ,  $p=.01$ ).

**Conclusion:** This study was able to demonstrate the predictive effect of a clinical OSPE regarding scores achieved on the state exam. Taking the limitations of this study into account, we are able to recommend using the OSPE testing format in the sixth semester during the clinical phase of dental study.

**Keywords:** OSPE, OSCE, predictive validity, reliability, state exams, length of study, clinical competency

Petkov Petko<sup>1</sup>  
Katja Knuth-Herzig<sup>2</sup>  
Sebastian Hoefler<sup>3</sup>  
Sebastian Stehle<sup>2</sup>  
Sonja Scherer<sup>2</sup>  
Björn Steffen<sup>4</sup>  
Stephan Scherzer<sup>5</sup>  
Falk Ochsendorf<sup>6</sup>  
Holger Horz<sup>2</sup>  
Robert Sader<sup>3</sup>  
Susanne  
Gerhardt-Szép<sup>1</sup>

- 1 Goethe-Universität, Carolinum Zahnärztliches Universitäts-Institut gGmbH, Poliklinik Zahnerhaltungskunde, Frankfurt am Main, Deutschland
- 2 Goethe-Universität, Institut für Psychologie & Interdisziplinäres Kolleg Hochschuldidaktik, Frankfurt am Main, Deutschland
- 3 Goethe-Universität, Klinik für Mund-, Kiefer- und Plastische Gesichtschirurgie, Frankfurt am Main, Deutschland
- 4 Goethe-Universität, Medizinische Klinik II, Hämatologie und Internistische Onkologie, Frankfurt am Main, Deutschland
- 5 Goethe-Universität, Zentrum für Innere Medizin, Frankfurt am Main, Deutschland
- 6 Goethe-Universität, Klinik für Dermatologie, Venerologie und Allergologie, Frankfurt am Main, Deutschland

# 1. Introduction and research question

Based on the German licensing rules and regulations for dentists (Approbationsordnung) [[https://www.gesetze-im-internet.de/\\_appro\\_2002/](https://www.gesetze-im-internet.de/_appro_2002/) cited 2015 October 22], dental education consists of a preclinical and clinical study phase, each lasting five semesters. Over the course of this, students acquire both theoretical knowledge and spend a considerable amount of time learning practical skills. During the preclinical phase basic knowledge of natural science and the scientific principles of medicine are taught, while during the later clinical phase, students directly apply the theoretical knowledge acquired in lectures and begin developing professional routines in the first clinical semester, initially using phantom patients (see figure 1) and then as part of multiple courses on treating real patients (see figure 2).

In the first clinical semester the main focus is on conservative dentistry. Students are taught in the simulation lab. To prepare and qualify students for the first course in treating patients, specific procedures are simulated and the required manual skills and techniques of restorative dentistry and endodontics are imparted through systematic training.

Acquisition of dental competencies during education requires close interlinking of teaching and testing, referred to as constructive alignment [1]. Consequently, the analysis of suitable testing formats has been at the center of research in dental education for years now [2]. A broad range of workplace-based assessments are used in dental education, especially in the clinical courses that center on treating patients. Testing formats are used which not only cover students' factual and practical knowledge, but also their practical skills [3].

Various internationally known testing methods exist to assess students' practical skills in complex situations, such as the Clinical Evaluation Exercise (CEX) [4], Mini-Clinical Evaluation Exercise (mCEX) [5], Entrustable Professional Activities (EPA) [6], Directly Observed Procedural Skills (DOPS) [6], and portfolios [7], all of which have not found wide use in dental medicine due to various problems such as insufficient content validity, questionable reliability, low acceptance and cumbersome implementation. However, dental education frequently relies on testing formats that reflect the "shows how" level of Miller's pyramid. These formats assess practical skills and communicative competencies, with specific reference here to the OSCE (Objective Structured Clinical Examination) [8], [9], OSPE (Objective Structured Practical Examination) [10] and the use of trained simulated or standardized patients, for instance as part of an OSCE [11], [12]. In the international literature the terms "OSCE" and "OSPE" are commonly used as synonyms. Since its introduction to the medical curriculum in 1975 [13], the OSCE has been successfully implemented around the world as an assessment tool in a wide variety of subjects [14], including dentistry exams since the 1990s [8], [9], [15]. In

addition to its wide acceptance among students, who feel the OSCE is a more just and less stressful testing format than traditional written and oral exams [16], it also strengthens skills in the area of clinical competence [9], [11], [17], fosters both communication skills [9] and learning, and contributes to a more accurate self-assessment by the participants [16], [17], [18].

During an OSCE students go through different stations at which specific practical skills (including individual steps of a treatment) are performed or a doctor-patient conversation is held (simulated patients). A special type of OSCE is the OSPE in which practical skills, knowledge, and interpretation of data are demonstrated in a non-clinical situation (e.g. simulation lab). In contrast to the OSCE, an OSPE provides an opportunity to assess entire work processes (all individual steps and the "finished product," for instance, a filling or an inlay in the field of conservative dentistry).

The dental program in Frankfurt am Main uses OSCEs (third clinical semester in the subjects conservative dentistry and oral and maxillofacial surgery), OSPEs (at the end of the first clinical semester in conservative dentistry), and simulated patients (in OSCEs). In the final stage of the degree program in dental medicine during the state examination, all students take not only theoretical but also practical tests in which everything must be performed independently on a patient (see figure 3).

Achieving the goal of best preparing students to treat patients requires critical evaluation of the concept underlying the existing rules and regulations in dental education governing the teaching of practical skills and abilities compared to theoretical knowledge of dental medicine. Early detection of deficiencies could be addressed and remedied individually by students and teachers over the course of study leading up to the state examination. However, it would be necessary to show evidence as to whether or not different testing formats correlate with each other, even if they each measure the same competencies (e.g. practical skills and abilities). When doing this, factors such as the sex of a student, length of study, score on the school-leaving exam and first section of the state medical exam should be examined for potential influence and predictive value. Accordingly, the primary aim of this study was to ascertain if, for the subject conservative dentistry, the final score on the OSPE in the sixth semester predicted the later practical exam score. In addition, it was sought to clarify whether or not correlations exist between the partial scores of the OSPE administered in the sixth semester and the corresponding practical scores on the state exam.

## 2. Material and Methods

### 2.1. Study timeline and setting

For this longitudinal retrospective study data was collected for two semesters (first and last clinical semesters). The period investigated was the time between the begin-



Figure 1: Treatment procedures on phantom patients during the first semester of clinical study



Figure 2: Treating patients during the state exam

ning of the 2009 summer semester and the end of the 2014 summer semester. People for whom information was available regarding all the listed variables were included in the analysis. In addition to dropping out between the OSPE and state exam ( $n=39$ , 14.61%), the reasons for exclusion from the calculations included missing data for scores on the Physikum (48 cases), scores on the school-leaving exam (33 cases), sex (six cases), or length of university study (219 cases). Another reason entailed OSPE data for students that was available for current semesters but with which no subsequent state exam scores could be matched at the time of our calculations. For students classified as drop-outs it is not known if they represent authentic instances of quitting university study

or if their scores were simply no longer adequately documented.

## 2.2. Student participants

Participants in the study were students enrolled in dental medicine at the Dental School (Carolinum) of the University of Frankfurt am Main, Germany. This study analyzed the available OSPE scores (retrospectively as of the 2009 summer semester) and the state examination (retrospectively as of the 2011/12 winter semester). Personal information was also evaluated, such as age, sex, school-leaving exam score, preliminary medical examination score, and length of university study, and students were

Task A: Filling	
1	Approximal surfaces and points of contact
2	Cavity base and walls
3	Axiopulpal (ap) line angles
4	Width and depth
5	Primary preparation
6	Smoothness of base filling and shape of ap line
7	Localization of base filling
8	Height of base filling
9	Base filling & secondary preparation
10	Margin sealing
11	Contact points
12	Occlusal contouring
13	Smoothness
14	Filling
15	OVERALL Filling score
Task B: Inlay	
16	External cavity walls
17	Internal cavity walls
18	Width and depth
19	Smoothness
20	Adjacent tooth
21	OVERALL Inlay score
22	FINAL OSPE SCORE

Figure 3: The individual items of the OSPE

assigned to cohorts. The sample population investigated here was comprised of 223 students, of which 141 were women and 82 men.

## Academic Setting

### 2.2.1. Sixth semester

At the end of the first clinical semester (sixth semester of specialized study) oral and practical assessments are given in the subject of conservative dentistry to measure knowledge gain. The semester-end, practical exam at the University of Frankfurt am Main's Dental School (Carolinum) takes the form of an OSPE, during which individual evaluations are recording using standard checklists. At Frankfurt am Main, three to four examiners are used for OSPEs. The assessment consists of two tasks – a dental filling and an inlay. Both tasks are divided into sub-steps. For the dental filling these are:

1. Primary preparation,
2. Bases/liners and secondary preparation,
3. Filling,
4. Filling overall.

The inlay station is divided into:

1. External cavity walls,
2. Internal cavity walls,

3. Width and depth,
4. Smoothness,
5. Adjacent tooth.

The individual items are presented in figure 3. They are evaluated using the traditional academic grading scale from 1 (very good) to 5 (deficient) and combined into a total score for each station. The overall evaluation of the filling consists of the individual scores for “primary preparation”, “bases/liners and secondary preparation”, and “filling”. At the end of the assessment a final score is calculated for the OSPE by combining the overall scores for the filling and inlay.

### 2.2.2. Semester of the state examination

As part of the state exam in dental medicine (11<sup>th</sup> semester of clinical study), the practical assessment in conservative dentistry lasts five days. Examinees are required to independently perform several procedures on patients, among them a root canal, preparation for an inlay and cementing the restoration into place, an anterior tooth filling, a posterior tooth filling, and making a diagnosis in two patients. The testing format used for the state exam covers procedural skills, as does the sixth-semester OSPE, but on the highest level of competency. The two items of overlap between the OSPE administered in the sixth-semester and the state exam – the filling and inlay

– are both covered to the same extent. Performance on the state exam is evaluated by two examiners using standardized checklists to assign scores. Filling and inlay stations relevant to this comparative study encompass the following sub-steps. For the filling station:

1. Primary preparation,
2. Rubber dam,
3. Base/liner and secondary preparation,
4. Matrices,
5. Filling before polishing,
6. Filling after polishing,
7. Filling overall.

For the inlay station:

1. Preparation (primary and secondary),
2. Impression,
3. Provisional inlay,
4. Inlay before placement,
5. Inlay after placement,
6. Inlay overall.

The exam checklist contains the same subcategories for the items filling and inlay as the sixth-semester OSPE checklist. The sub-step “rubber dam” is covered in the OSPE evaluation of bases/liners and secondary preparation; matrices and the steps involving the filling before and after polishing are part of the OSPE section covering the filling. The five sub-steps for the OSPE inlay station are part of “preparation” in the state exam. Steps two through five (impression to inlay after cementing) are not covered by the OSPE because an indirect restoration is not part of that assessment. This would take several days since these fillings would have to be made in a dental lab and thus cannot be realized in a three-hour OSPE. The individual steps in the state exam are also evaluated using the traditional grading scale of 1 (very good) to 5 (deficient).

## 2.3. Analytical methods

Statistical analysis was performed in cooperation with the Department of Educational Psychology (Frankfurt am Main). Analysis was carried out using SPSS Statistics 21 (IBM Corporation, New York) and calculated Spearman rank correlation, partial correlation, Pearson’s correlation coefficient and (multiple) linear regressions.

## 3. Results

### 3.1. General results

#### 3.1.1. Reliability of the OSPE

Based on all partial scores, The OSPE demonstrates overall a high internal consistency with a Cronbach’s  $\alpha$  of .87. The two items for filling (Cronbach’s  $\alpha$ =.84) and inlay (Cronbach’s  $\alpha$ =.87) also reflect a good internal consistency.

#### 3.1.2. Prediction of state exam scores based on OSPE performance

In the student population investigated here, the final OSPE score correlated to a significantly positive degree with the overall score on the practical section of the state exam (Spearman rank correlation,  $r$ =.14,  $p$ =.03). It can be observed that higher OSPE scores go along with higher practical exam scores. Students earned a mean of 3.79 (standard deviation=0.81) on the OSPE and were able to show a mean improvement of approximately two grade levels ( $M$ =1.87,  $SD$ =.38).

Of the variables investigated, apart from the OSPE score, only the length of study (Spearman rank correlation,  $r$ =.20,  $p$ =.003) correlated significantly with the exam score, showing that a higher exam score paired with a shorter length of study, and a longer length of study with a lower exam score.

The calculation of a multiple regression with stepwise addition of the OSPE score and the length of study showed that both contribute significantly to predicting the state exam score ( $\beta$ =.16,  $p$ =.02,  $R^2$ =.08).

#### 3.1.3. Correlation between partial scores on the OSPE and partial scores on the state exam (filling and inlay)

The OSPE scores for filling and inlay were also correlated with the score for the corresponding items on the practical state exam (see table 1). These two items were selected because they allow relatively direct comparison due to extensive overlap of content and suggest a link between scores. It was possible to determine, while examining the cohort as a possible influential factor, that the OSPE and state exam scores for the filling did not correlate significantly with each other ( $r$ =.12,  $p$ =.07). The scores assigned for inlays, in contrast, showed such a correlation ( $r$ =.13,  $p$ =.05).

Within an exam or assessment, however, these two scores correlated quite highly with each other, meaning that whoever did well (or poorly) on a section of a particular test also did as well (or poorly) on the other section. This is reflected in the good reliability of the OSPE overall assessments (see table 1).

### 3.2. Gender differences

#### 3.2.1. Scores and length of study

In the sample of 141 female and 82 male students, no statistically significant differences could be determined within the OSPE regarding sex and the partial scores for filling and inlay (indirect restoration) or the final score. Furthermore, in respect to the partial scores and final scores on the state exam there were no gender differences. The same is true for length of study. Gender differences could only be detected for the scores on the school-leaving exam and the Physikum (preliminary exam in dental medicine). The school-leaving exam scores showed that female students ( $M$ =2.12) had significantly higher

**Table 1: Correlation of the total and partial scores for the filling and inlay stations in both test settings (.\*\*. The correlation is significant on the level of .01 (two-sided).**

	Exam Inlay	Exam Filling	OSPE Filling
Exam Filling	.246** (r < .000)	/	/
OSPE Filling	.042 (r = .538)	.123 (r = .067)	/
OSPE Inlay	.133* (r = .048)	.162* (r = .016)	.360** (r = .000)

scores ( $F=6.09$ ,  $p=.01$ ,  $\eta^2=.03$ ) than male students ( $M=2.31$ ). However, no correlations between the school-leaving exam scores could be seen with the OSPE or state exam, and consequently success on the OSPE or practical state exam cannot be predicted by the school-leaving exam score. This could be explained by the use of the school-leaving exam score as a selective criterion when admitting students to the degree program, something that definitively limits its variance in the sample population. In the Physikum it was seen that male students with a mean value of 2.24 performed significantly better ( $F=4.15$ ,  $p=.04$ ,  $\eta^2=.02$ ) than female students ( $M=2.40$ ). In male students this score also correlated significantly ( $r=.24$ ,  $p=.03$ ) with the OSPE, with a high OSPE score accompanying a high Physikum score. This was not the case for female students ( $r=.12$ ,  $p=.17$ ). In contrast, no significant correlations between Physikum score and state exam score could be determined.

### 3.2.2. Using the OSPE to predict the state exam score

If the correlation of the OSPE is considered separately according to sex, it is seen that for female students a higher correlation ( $r=.195$ ,  $p=.02$ ) can be found between the OSPE and the state exam score than is the case for male students ( $r=.137$ ,  $p=.22$ ). Likewise, a high correlation ( $r=.27$ ,  $p=.007$ ) exists between length of study and exam score only for female students if calculations are made separately according to sex. The calculation of a separate multiple regression analysis with stepwise addition of these two variables revealed that only for female students did they contribute significantly to predicting the exam score ( $\beta=.18$ ,  $p=.03$ ,  $R^2=.10$ ). This was lower for male students ( $p=.16$ ,  $R^2=.05$ ).

## 4. Discussion

Well-known internationally, the testing format applied by the OSCE demonstrates reliability values between Cronbach's  $\alpha$  .40 and .91 [19]. This can be shown for both the purely medical field and dental medicine [10], [18]. The OSPE represents a variation of the OSCE that is focused more on practical skills [10]. Both terms are often used in the literature to mean the same thing [10]. Studies on the predictive validity of the OSCE/OSPE in medicine and dental medicine [8], [15], [17], [18], [20] show that student performance on OSCEs reflects their

future clinical competency. Our study confirms a high reliability and a predictive validity of success on the practical state exam in conservative dentistry for a specific format of the OSPE, thus corroborating the studies cited above. In contrast, a study of dental medicine [21] found no significant correlation between OSCE and state exam. A possible reason could be that the overall score for that assessment (both practical and oral assessments) was investigated, while this study concentrates on direct comparison of the practical skills. No other publications are known in which the OSCE/OSPE and state exam have been investigated in terms of scores on identical assessment subsections. Within this context, this study presents a new perspective on this issue. It is surprising that the correlation between OSPE and the practical state exam is much clearer for women, something that raises a series of interesting questions. However, the limitation of this study must be kept in mind. Based on the available data, the potential influence of the task to be performed and rater cannot be analyzed using the methods applied here. Moreover, it was possible to determine that the scores for inlays correlated significantly with each other for the OSPE and state exam, but this was not the case for fillings.

The reason for this could lie in the different test settings (OSPE versus state exam). However, it remains that whether the assessment is conducted in a clinical setting on phantom patients (OSPE) or on real patients (state exam), the tasks of filling and inlay obviously require a very similar level of learning transfer and manual dexterity within a particular assessment. Within an assessment these two partial scores correlate with each other. When different assessments are compared, the requirements for inlays appear to be more similar than those for the filling. It can be supposed that the ability to visualize spatially, a skill that generally plays an important role in exercising practical skills, exerts an influence across assessment settings.

The general increase in practical competencies from the OSPE (mean score 3.79) to the state exam (mean 1.87) can, in our view, be traced to a positive training effect. These observations concur with the results of the study by Sloan [20].

Furthermore, it has been shown that female students achieve significantly higher scores on the school-leaving exam (Abitur), although the range of scores for the sample in this study is severely limited as a result of its use as a criterion for university admission. Similar results are seen

in an analysis of the data for over 126,000 students who passed the school-leaving exam thereby attaining the formal qualification for university study [http://www.it.nrw.de cited 2015 October 22]. However, it must be noted that the score on the school-leaving exam does not show any correlation with OSPE scores or those for the practical section of the state exam.

In comparison, for male students the score on the Physikum was higher and a correlation was seen between OSPE score and Physikum score showing that a stronger performance on the OSPE is accompanied by a stronger performance on the Physikum. This correlation could not be found in female students. No comparable studies have been published that could be cited in connection with these results. The fact is that, in contrast to the OSPE score which measures practical skill, the Physikum score involves a conglomerate of four oral assessments (anatomy, physiology, biochemistry, and prothodontics). The grade in prosthodontics includes a practical evaluation that could not be retrospectively filtered out in our study setting. For this reason it is not possible to identify which common factor or factors influence these two scores achieved by male students. Moreover, the analyses in this study revealed a correlation between study length and performance on the practical exam. A longer period of study correlates significantly with a lower exam score. Identical results were confirmed by a study in the field of dental education by Eberhard [21]. A longer length of study could compromise student motivation, or the sense of belonging to a particular semester cohort could suffer so much that any positive effects connected with belonging could be lost. In addition, the longer period of study could be an indication of external influences that affect not only the final grade but also the length of time spent studying at the university. It is conceivable that a student's financial circumstances or family situation could assert this kind of influence. It would make sense that the influence is only present in female students if the sample population is classified by sex. It has already been shown that female students are more affected by multiple burdens such as pursuing academic study and meeting family responsibilities [22]. There are several indications in the results seen here—the low and moderately high correlations and small effect sizes—that other influential factors are missing from the model. This is not surprising since this study only takes performance measures and similar variables into account. Comprehensive personal profiles have not been available to date. Also, the data regarding the characteristics of the (testing) situation are incomplete. The reasons for these two limitations are found in the retrospective collection of the data. Another aspect which must be viewed critically is the fact that this study investigated only complete cases which can limit the validity of conclusions drawn from statistical analyses.

## 5. Conclusions

In light of the results seen here, it is possible to assert that the testing format of the OSPE partially predicts success on the practical section of the state exam covering conservative dentistry even if the results show, as expected, additional influences, such as length of study, and indicate the possible presence of other factors not visible in the data at hand. In future research, it would be worth the effort to monitor such factors over the course of study.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Biggs J. Enhancing teaching through constructive alignment. *High Educ.* 1996;32(3):347-364. DOI: 10.1007/BF00138871
2. Abou-Rass MA. Clinical evaluation instrument in endodontics. *J Dent Educ.* 1973;37(9):22-36.
3. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9 Suppl):S63-67. DOI: 10.1097/00001888-199009000-00045
4. Thornton S. A literature review of the long case and its variants as a method of assessment. *Educ Med J.* 2012;4(1).
5. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med.* 2003;138(6):476-481. DOI: 10.7326/0003-4819-138-6-200303180-00012
6. Hauer KE, Kohlwes J, Cornett P, Hollander H, Cate ten O, Ranji SR. Identifying entrustable professional activities in internal medicine training. *J Grad Med Educ.* 2013;5(1):54-59. DOI: 10.4300/JGME-D-12-00060.1
7. Gadbury-Amyot CC, McCracken MS, Woldt JL, Brennan RL. Validity and reliability of portfolio assessment of student competence in two dental school populations: a four-year study. *J Dent Educ.* 2014;78(5):657-667.
8. Graham R, Bitzer LAZ, Anderson OR. Reliability and Predictive Validity of a Comprehensive Preclinical OSCE in Dental Education. *J Dent Educ.* 2013;77(2):161-167.
9. Ratzmann A, Wiesmann U, Kordaß B. Integration of an Objective Structured Clinical Examination (OSCE) into the dental preliminary exams. *GMS Z Med Ausbild.* 2012;29(1):Doc09. DOI: 10.3205/zma000779
10. Schmitt L. Reliabilität der OSPE (Objective Structured Practical Examination) in Abhängigkeit von der Prüferanzahl und den Bewertungskriterien im ersten klinischen Semester des Faches Zahnerhaltungskunde, Promotion. Frankfurt: Goethe Universität Frankfurt am Main; 2014.
11. Hallberg J. Quantitative Analyse der OSCEs eines Phantomkurses der Zahnerhaltung und Parodontologie, Promotion. Hannover: Medizinische Hochschule Hannover; 2012.
12. Ortwein H, Fröhmel A, Burger W. Einsatz von Simulationspatienten als Lehr-, Lern und Prüfungsform. *Psychother Psych Med.* 2006;56:23-29. DOI: 10.1055/s-2005-867058
13. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *BMJ.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447

14. Nayak V, Bairy KL, Adiga S, Shenoy S, Chogtu BM, Amberkar M, Kumari KM. OSPE in Pharmacology: Comparison with the conventional Method and Students' Perspective Towards OSPE. *Br Biomed Bull.* 2014;2(1):218-222.
15. Martin IG, Jolly B. Predictive validity and estimated cut score of an objective structured clinical examination (OSCE) used as an assessment of clinical skills at the end of the first clinical year. *Med Educ.* 2002;36(5):418-425. DOI: 10.1046/j.1365-2923.2002.01207.x
16. Smith LJ, Price DA, Houston IB. Objective structured clinical examination compared with other forms of student assessment. *Arch Dis Child.* 1984;59:1173-1176. DOI: 10.1136/adc.59.12.1173
17. Schoonheim-Klein ME, Habets LL, Aartman IH, van der Vleuten CP, Hoogstraten 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(4):226-235. DOI: 10.1111/j.1600-0579.2006.00421.x
18. Brown G, Manogue M, Martin M. The validity and reliability of an OSCE in dentistry. *Eur J Dent Educ.* 1999;3(3):117-125. DOI: 10.1111/j.1600-0579.1999.tb00077.x
19. Cohen R, Reznick RK, Taylor BR, Provan J, Rothman A. Reliability and validity of the objective structured clinical examination in assessing surgical residents. *Am J Surg.* 1990;160(3):302-305. DOI: 10.1016/S0002-9610(06)80029-2
20. Sloan DA, Donnelly MB, Schwartz RW, Felts JL, Blue AV, Strodel WE. The use of objective structured clinical examination (OSCE) for evaluation and instruction in graduate medical education. *J Surg Res.* 1996;63(1):225-230. DOI: 10.1006/jsre.1996.0252
21. Eberhard L. Der erste Eindruck zählt? – Prädiktive Validität einer vorklinischen OSCE. Vortrag auf der AKWLZ-Tagung in Ulm. Ulm: Universität Ulm; 2014.
22. Middendorf E, Apolinarski B, Poskowsky J, Kandulla M, Netz N. Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2012. 20. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch das HIS-Hochschul-Informationen-System. Bonn: Bundesministerium für Bildung und Forschung; 2013.

**Corresponding author:**

PD Dr. med. dent. Susanne Gerhardt-Szép, MME  
Goethe-University, Dental School (Carolinum), Department  
of Operative Dentistry, D-60596 Frankfurt am Main,  
Germany, Phone: +49 (0)69/6301-7505, Fax: +49  
(0)69/6301-3841  
s.szep@em.uni-frankfurt.de

**Please cite as**

*Petko P, Knuth-Herzig K, Hoefer S, Stehle S, Scherer S, Steffen B, Scherzer S, Ochsendorf F, Horz H, Sader R, Gerhardt-Szép S. The reliability and predictive validity of a sixth-semester OSPE in conservative dentistry regarding performance on the state examination. GMS J Med Educ.* 2017;34(1):Doc10.  
DOI: 10.3205/zma001087, URN: urn:nbn:de:0183-zma0010873

**This article is freely available from**

<http://www.egms.de/en/journals/zma/2017-34/zma001087.shtml>

**Received:** 2015-10-23

**Revised:** 2016-10-20

**Accepted:** 2016-11-09

**Published:** 2017-02-15

**Copyright**

©2017 Petko et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.



# Reliabilität und prädiktive Validität einer OSPE im 6. Semester im Fach Zahnerhaltungskunde im Hinblick auf die Staatsexamensnote

## Zusammenfassung

**Einleitung:** Ziel dieser Studie war es zu evaluieren, ob das Prüfungsformat einer OSPE (Objective Structured Practical Examination) durchgeführt im Fach Zahnerhaltungskunde (6. Fachsemester) den Studienerfolg im praktischen Teil des Staatsexamens (11. Fachsemester) im selben Fach prädiziert. Ferner sollte unter Berücksichtigung allgemeiner Angaben der StudienteilnehmerInnen (Abiturnote, Physikumsnote, Studiendauer, Kohorte und Geschlecht) analysiert werden, ob bezüglich der Gesamt- sowie Teilnoten der OSPE und der adäquaten Staatsexamensprüfung Zusammenhänge oder Unterschiede bestehen.

**Methoden:** Im Rahmen dieser longitudinalen, retrospektiven Studie wurden für einen Zeitraum von 11 Semestern prüfungsbezogene Daten von Studierenden (N=223) des Fachbereichs Zahnmedizin in Frankfurt am Main erhoben und untersucht. Für die statistische Auswertung der Daten wurden Spearman Rangkorrelationen, Partialkorrelationen, Korrelationskoeffizienten nach Pearson, und Multiple Regressionen (SPSS Statistics 21, IBM Corporation, New York) berechnet.

**Ergebnisse:** Die Ergebnisse zeigen, dass OSPE (Cronbachs  $\alpha=.87$ ) mit dem Erfolg im praktischen Teil des Staatsexamens im Fach Zahnerhaltungskunde korreliert ( $p=.01$ ,  $r=.17$ ). Als eine weitere signifikante Korrelation mit der Examensleistung erwies sich die Dauer des Studiums ( $p=.001$ ,  $r=.23$ ). Gemeinsam leisten diese beiden Variablen einen signifikanten Beitrag zur Vorhersage der Examensnote ( $p=.001$ ,  $R^2=.076$ ). Das zeigte sich im größeren Umfang bei weiblichen Studierenden. Zudem wurde festgestellt, dass diese bessere Abiturnoten als männliche Studierende aufweisen ( $F=6.09$ ,  $p=.01$ ,  $\eta^2=.027$ ) und dass es lediglich bei männlichen Studierenden eine signifikante Korrelation zwischen der Physikumsnote (Zahnärztliche Vorprüfung) und der OSPE-Benotung gab ( $r=.17$ ,  $p=.01$ ).

**Schlussfolgerung:** In der vorliegenden Untersuchung konnte der prädiktive Effekt einer klinischen OSPE auf die Prüfungsleistung im Staatsexamen gezeigt werden. Unter Berücksichtigung der Limitation der Studie empfiehlt sich aus unserer Sicht die Durchführung eines solchen Prüfungsformats im Rahmen des klinischen Studienabschnitts im 6. Semester im Fach Zahnmedizin.

**Schlüsselwörter:** OSPE, OSCE, Prädiktive Validität, Reliabilität, Staatsexamen, Studiendauer, klinische Kompetenz

Petkov Petko<sup>1</sup>  
Katja Knuth-Herzig<sup>2</sup>  
Sebastian Hoefler<sup>3</sup>  
Sebastian Stehle<sup>2</sup>  
Sonja Scherer<sup>2</sup>  
Björn Steffen<sup>4</sup>  
Stephan Scherzer<sup>5</sup>  
Falk Ochsendorf<sup>6</sup>  
Holger Horz<sup>2</sup>  
Robert Sader<sup>3</sup>  
Susanne  
Gerhardt-Szép<sup>1</sup>

1 Goethe-Universität,  
Carolinum Zahnärztliches  
Universitäts-Institut gGmbH,  
Poliklinik  
Zahnerhaltungskunde,  
Frankfurt am Main,  
Deutschland

2 Goethe-Universität, Institut  
für Psychologie &  
Interdisziplinäres Kolleg  
Hochschuldidaktik, Frankfurt  
am Main, Deutschland

3 Goethe-Universität, Klinik für  
Mund-, Kiefer- und Plastische  
Gesichtschirurgie, Frankfurt  
am Main, Deutschland

4 Goethe-Universität,  
Medizinische Klinik II,  
Hämatologie und  
Internistische Onkologie,  
Frankfurt am Main,  
Deutschland

5 Goethe-Universität, Zentrum  
für Innere Medizin, Frankfurt  
am Main, Deutschland

6 Goethe-Universität, Klinik für  
Dermatologie, Venerologie  
und Allergologie, Frankfurt  
am Main, Deutschland

# 1. Einleitung und Problemstellung

Geregelt durch die Approbationsordnung für Zahnärzte [https://www.gesetze-im-internet.de/\_appro\_2002/ zitiert am 22.10.2015], besteht das Zahnmedizinstudium aus einem vorklinischen und einem klinischen Studienabschnitt mit einer Dauer von jeweils fünf Semestern. Dabei spielen sowohl der theoretische Wissenserwerb als auch in beträchtlichem Ausmaß das Erlernen praktischer Fertigkeiten eine Rolle. Im vorklinischen Ausbildungsabschnitt werden naturwissenschaftliche Grundlagen und die wissenschaftlichen Prinzipien der Medizin vermittelt, während im klinischen Abschnitt parallel zum theoretischen Wissenserwerb durch Vorlesungen auch die direkte Wissensanwendung und das Erlangen therapeutischer Routine im ersten klinischen Semester zunächst am Phantompatienten (siehe Abbildung 1) und darauffolgend im Rahmen von mehreren Patientenbehandlungskursen erfolgt (siehe Abbildung 2).

Im ersten klinischen Semester stellt das Fach Zahnerhaltungskunde den Schwerpunkt dar. Hier werden die Studierenden im Simulationslabor ausgebildet. Zur Qualifizierung der Studierenden für den ersten Behandlungskurs am Patienten werden patientennahe Behandlungsabläufe simuliert und durch ein systematisches Training die erforderlichen grundlegenden manuellen Fertigkeiten und Techniken der restaurativen Zahnheilkunde und Endodontie vermittelt.

Das Erlangen zahnmedizinischer Kompetenzen während der Ausbildung erfordert eine enge Verknüpfung von Prüfungen und Lehre, „constructive alignment“ [1]. Aus diesem Grund steht die Analyse geeigneter Prüfungsformate seit vielen Jahren im Fokus der zahnmedizinischen Lehrforschung [2]. Im Zahnmedizinstudium werden vornehmlich die verschiedensten Formen des arbeitsplatzbasierten Prüfens in den klinischen Patientenbehandlungskursen praktiziert. Hier werden Prüfungsformate angewendet, die nicht nur Fakten- und Handlungswissen erfassen, sondern auch die praktischen Kompetenzen der Studierenden [3].

Es existieren verschiedene, international beschriebene Prüfungsmethoden zur Überprüfung der praktischen Kompetenz von Studierenden in komplexen Situationen wie z.B. Clinical Evaluation Exercise (CEX) [4], Mini-Clinical Evaluation Exercise (mCEX) [5], Entrustable Professional Activities (EPA) [6], Directly Observed Procedural Skills (DOPS) [6], Portfolio [7], die allesamt aufgrund verschiedener Faktoren wie unzureichender Inhaltsvalidität, fraglicher Reliabilität, geringer Akzeptanz oder umständlicher Durchführbarkeit bislang nur begrenzte Anwendung in der Zahnmedizin finden. Häufig hingegen werden in der Zahnmedizin Prüfungsformen angewendet, die die Millersche Ebene „zeigt wie“ abbilden. Diese Settings bieten sich zur Überprüfung praktischer Fertigkeiten und kommunikativer Kompetenzen an, wobei in diesem Zusammenhang vor allem OSCEs (Objective Structured Clinical Examinations) [8], [9], OSPEs (Objective Structured Practical Examinations) [10] und der Einsatz von geschul- ten Simulations- bzw. standardisierten Patienten, z.B. im

Rahmen einer OSCE, erwähnt werden [11], [12]. In der internationalen Literatur werden die Begriffe OSCE und OSPE meist synonym verwendet. Das Prüfungsformat OSCE wird seit seiner Einführung im Jahr 1975 [13] im Fach Medizin weltweit in vielen verschiedenen Fachgebieten [14], sowie seit den 1990er Jahren auch im Rahmen zahnmedizinischer Prüfungen [8], [9], [15] erfolgreich als Prüfungsinstrument eingesetzt. Neben seiner breiten Akzeptanz unter den Studierenden, die OSCE im Vergleich zu traditionellen schriftlichen und mündlichen Prüfungen als gerechtere und stressfreiere Prüfungsform empfinden [16], verstärkt diese Prüfungsform zusätzlich die Fähigkeiten im Bereich der klinischen Kompetenz [9], [11], [17], fördert die kommunikativen Fähigkeiten [9], ebenso wie das Lernen selbst und trägt zu einer besseren Selbsteinschätzung der Teilnehmer bei [16], [17], [18]. Im Verlauf einer OSCE durchlaufen die Studierenden verschiedene Stationen, bei denen jeweils praktische Tätigkeiten (auch Teilabschnitte einer Behandlung) oder Patientengespräche (Simulationspatienten) durchgeführt werden. Eine Spezialform von OSCEs stellen OSPEs dar, bei denen praktische Fertigkeiten, Wissen und Interpretation von Daten in einer nichtklinischen Situation (z.B. im Simulationslabor) erfolgen. Im Unterschied zum OSCE können im Rahmen einer OSPE ganze Arbeitsprozesse geprüft werden (alle Einzelschritte und das fertige „Endprodukt“, z.B. im Bereich der Zahnerhaltungskunde eine Füllung oder ein Inlay).

In Frankfurt am Main werden OSCEs (im 3. klinischen Semester in den Fächern Zahnersatzkunde sowie Mund-, Kiefer-, Gesichtschirurgie), OSPEs (am Ende des ersten klinischen Semesters im Fach Zahnerhaltungskunde) und Simulationspatienten (im Rahmen von OSCEs) eingesetzt. Zum Abschluss des zahnmedizinischen Studiums, während des Staatsexamens, werden von allen Studierenden neben theoretischen auch praktische Prüfungen abgelegt, wobei hier alle Leistungen ausschließlich selbstständig am Patienten erbracht werden (siehe Abbildung 3).



**Abbildung 1: Behandlung am Phantompatienten im ersten klinischen Semester**



**Abbildung 2: Behandlung am Patienten im Staatsexamen**

	<b>Aufgabe A: Füllung</b>
1	Approximale Ausserkontaktstellung
2	Boden und Wände
3	Winkel pulpaaxiale Wand
4	Breite und Tiefe
5	Primärpräparation
6	UF glatt und pulpaaxiale Ausformung
7	UF-Lokalisation
8	UF-Höhe
9	Unterfüllung & Sekundärpräparation
10	Randdichte
11	Kontaktpunkte
12	Okklusale Gestaltung
13	Glätte
14	Füllung
15	Füllung GESAMTNOTE
	<b>Aufgabe B: Inlay</b>
16	Kavitätenausenränder
17	Kavitäteninnenränder
18	Breite und Tiefe
19	Glätte
20	Nachbarzahn
21	Inlay GESAMTNOTE
22	GESAMTNOTE OSPE

Abbildung 3: Die einzelnen Items des OSPE

Das Ziel, eine bestmögliche Vorbereitung der Studierenden auf die Behandlung der Patienten zu gewährleisten, erfordert eine kritische Überprüfung des Konzepts der bestehenden Regelungen zur Ausbildung praktischer Fertigkeiten und Fähigkeiten im Vergleich zu theoretischen Kenntnissen in der Zahnmedizin. Frühzeitig erkannte Kompetenzmängel könnten im Verlauf bis zum Staatsexamen individuell von den Studierenden zusammen mit den Lehrenden aufgearbeitet werden. Dafür müsste jedoch Evidenz darüber erhoben werden, ob verschiedene Prüfungsformate, selbst wenn sie jeweils die identischen Kompetenzen (beispielsweise die praktischen Fertigkeiten und Fähigkeiten) abbilden, miteinander korrelieren. Dabei sollten auch Faktoren, wie beispielsweise das Geschlecht der Studierenden, die Studiendauer, die Abiturnote oder die Physikumsnote, auf einen möglichen Einfluss und prädiktiven Wert hin überprüft werden. Das Hauptziel dieser Studie war dementsprechend, zu überprüfen, ob die Gesamtnote des OSPE im 6. Semester die praktische Examensnote im Fach Zahnerhaltungskunde prädiziert. Zudem sollte geklärt werden, ob es Korrelation zwischen den einzelnen Teilnoten des OSPE im 6. Semester und den korrespondierenden praktischen Teilnoten im Examen gibt.

## 2. Material und Methoden

### 2.1. Untersuchungszeitraum und Setting

Für diese longitudinale retrospektive Studie wurden die Daten von insgesamt zwei Semestern (erstes und letztes klinisches Semester) erhoben. Der untersuchte Zeitraum lag zwischen Beginn des Sommersemesters 2009 und

Ende des Sommersemesters 2014. Es wurden Personen in die Berechnungen aufgenommen, von denen zu allen genannten Variablen Daten vorlagen. Neben einem Dropout zwischen den beiden untersuchten Prüfungen (n=39, 14,61%), waren auch das Fehlen der Angaben zur Physikumsnote (48 Fälle), Abiturnote (33 Fälle), Geschlecht (6 Fälle) oder Studiendauer (219 Fälle) ein Grund für den Ausschluss der Person aus den Berechnungen. Zudem lagen OSPE Daten von Studierenden aus aktuellen Semestern vor, zu denen es naturgemäß zum Zeitpunkt der Berechnungen noch keine Examensnote geben konnte. Bei den Personen, die als Dropout gewertet wurden liegen zudem keine Informationen vor, ob es sich um einen echten Dropout handelt oder lediglich Prüfungsnoten nicht mehr nachvollziehbar waren.

### 2.2. Studienteilnehmer

StudienteilnehmerInnen waren Lernende des Fachs Zahnmedizin des Carolinum Zahnärztliches Universitäts-Institut gGmbH Frankfurt am Main. Die Untersuchung erfolgt durch Analyse der vorliegenden Prüfungsergebnisse des OSPE (retrospektiv ab SS 2009), und des Staatsexamens (retrospektiv ab WS 2011/12). Außerdem wurden persönliche Angaben der Teilnehmer wie Alter, Geschlecht, Abiturnote, Physikumsnote und Studiendauer evaluiert sowie eine Zuordnung zu Kohorten gebildet. Die untersuchte Stichprobe bestand aus 223 Studierenden, davon 141 Frauen und 82 Männer.

### Bedingungen vor Ort

#### 2.2.1. Sechstes Semester

Am Ende des ersten klinischen Semesters (6. Fachsemester) werden im Fach Zahnerhaltungskunde die erworbenen Kenntnisse durch mündliche und praktische Prüfungen erfasst. Die praktische Semesterabschlussprüfung im Zahnärztlichen Universitäts-Institut Carolinum in Frankfurt am Main wird in Form einer OSPE durchgeführt. Dort werden Einzelbewertungen durch standardisierte Checklisten erhoben. Bei den OSPEs in Frankfurt am Main wurden drei bis vier PrüferInnen eingesetzt. Die Prüfung besteht aus zwei Aufgaben – Füllungstherapie und Inlay. Beide werden in Untereinheiten unterteilt. Zum Item Füllung sind dies:

1. Primärpräparation,
2. Unterfüllung und Sekundärpräparation,
3. Füllung und
4. Füllung gesamt.

Die Station Inlay wird unterteilt in

1. Kavitätenaußenränder
2. Kavitäteninnenränder
3. Breite und Tiefe
4. Glätte
5. Nachbarzahn

Die einzelnen Teilbewertungen (Items) sind in Abbildung 3 dargestellt. Sie werden jeweils mit Schulnoten zwischen 1 (sehr gut) und 5 (ungenügend) bewertet und zu einer Note für die entsprechende Station zusammengefasst. Die Bewertung „Füllung gesamt“ setzt sich aus den Einzelnoten „Primärpräparation“, „Unterfüllung/Sekundärpräparation“ und „Füllung“ zusammen. Am Ende der Prüfung wird aus den Benotungen „Füllung gesamt“ und „Inlay gesamt“ eine OSPE-Gesamtnote gebildet.

### 2.2.2. Staatsexamenssemester

Im Rahmen des zahnmedizinischen Staatsexamens (11. klinisches Fachsemester) dauert die praktische Prüfung im Fach Zahnerhaltungskunde fünf Tage. Die Prüfungskandidaten müssen mehrere Leistungen selbstständig am Patienten durchführen. Dazu gehören eine Wurzelkanalbehandlung, eine Präparation zur Aufnahme eines Inlays und die Eingliederung der fertigen Restauration, eine Frontzahnfüllung, eine Seitenzahnfüllung und zwei zahnmedizinische Patientenbefunde. Das Prüfungsformat im Staatsexamen testet ebenso wie die OSPE im 6. Semester die prozeduralen Fertigkeiten, allerdings auf dem höchsten Kompetenzniveau. Die mit OSPE im 6. Semester korrespondierenden Leistungen Füllung und Inlay sind hier im selben Umfang erhalten. Die Bewertung im Staatsexamen erfolgt durch zwei PrüferInnen und für die Notenvergabe werden ebenfalls standardisierte Checklisten eingesetzt. Die für unsere Vergleichsuntersuchung relevanten Stationen Füllung und Inlay beinhalten nachfolgend aufgeführte Teilschritte. Für Füllung:

1. Primärpräparation,
2. Kofferdam,
3. Unterfüllung und Sekundärpräparation,
4. Matrize,
5. Füllung vor Politur,
6. Füllung nach Politur,
7. Füllung gesamt.

für die Station Inlay:

1. Präparation (Primär und Sekundär),
2. Abformung,
3. Provisorium,
4. Inlay vor Eingliederung,
5. Inlay nach Eingliederung,
6. Inlay gesamt.

Für die Items Füllung und Inlay beinhaltet die Examenscheckliste mitunter dieselben Unterkategorien wie die OSPE-Checkliste im 6. Semester. Der Teilschritt „Kofferdam“ im Examen gehört in die Bewertung „Unterfüllung und Sekundärpräparation“ im OSPE, die „Matrize“ und die Schritte „Füllung vor bzw. nach der Politur“ sind Teil des Abschnittes „Füllung“ im OSPE. Die fünf Teilschritte der Station „Inlay“ im OSPE werden im Examen unter „Präparation“ subsummiert. Schritte zwei bis fünf (Abformung bis Inlay nach Eingliederung) sind im OSPE nicht abgebildet, weil im Rahmen dieser Prüfung keine indirekte Restauration angefertigt wird. Dies würde mehrere Tage

dauern, da sie in einem zahntechnischen Labor angefertigt wird und somit nicht Inhalt des dreistündigen OSPE zu realisieren ist. Die einzelnen Schritte werden auch im Examen jeweils mit Schulnoten zwischen 1 (sehr gut) und 5 (ungenügend) bewertet.

## 2.3. Auswertungsmethoden

Die statistische Analyse wurde in Zusammenarbeit mit dem Fachbereich Pädagogische Psychologie (Frankfurt am Main), durchgeführt. Alle Auswertungen wurden mit dem Programm SPSS Statistics 21 (IBM Corporation, New York) unter Berechnung von Spearmans Rangkorrelation, Partieller Korrelation, Korrelationskoeffizient nach Pearson und (Multiplen) Linearen Regressionen erstellt.

## 3. Ergebnisse

### 3.1. Allgemeine Ergebnisse

#### 3.1.1. Reliabilität des OSPE

Der OSPE zeigte, berechnet über alle Teilbewertungen insgesamt eine hohe interne Konsistenz von Cronbachs  $\alpha=.87$ . Die beiden Teilaufgaben für Füllung (Cronbachs  $\alpha=.84$ ) und Inlay (Cronbachs  $\alpha=.87$ ) erreichten ebenfalls eine gute interne Konsistenz.

#### 3.1.2. Prädiktion der Staatsexamensnoten durch OSPE

Bei der beschriebenen Stichprobe von Studierenden korrelierte die OSPE-Gesamtnoten signifikant positiv mit der praktischen Examensgesamtnote (Spearmans Rangkorrelation,  $r=.14$ ,  $p=.03$ ). Es zeigte sich, dass leistungsstärkere OSPE-Noten mit leistungsstärkeren praktischen Examensnoten einhergehen. Im Mittel erhielten die Studierenden im OSPE eine Bewertung von 3.79 (Standardabweichung/SD=.81) und konnten im Staatsexamen eine mittlere Verbesserung um etwa 2 Notenstufen ( $M=1.87$ ,  $SD=.38$ ) erreichen.

Neben der OSPE-Note korrelierte von den erfassten Variablen lediglich die Studiendauer (Spearmans Rangkorrelation,  $r=.20$ ,  $p=.003$ ) signifikant mit der Examensnote, wobei eine leistungsstärkere Examensnote mit einer kürzeren Studiendauer einherging, eine längere Studiendauer mit einer schlechteren Examensnote.

Die Berechnung einer multiplen Regression mit schrittweisem Hinzufügen der OSPE Note und der Studiendauer ergab, dass beide einen signifikanten Beitrag zur Vorhersage der Examensnote leisten ( $\beta=.16$ ,  $p=.02$ ,  $R^2=.08$ ).

#### 3.1.3. Zusammenhang zwischen Teilnoten des OSPE und Teilnoten der Staatsexamensprüfung (Füllung und Inlay)

Zusätzlich wurden die OSPE-Noten für Inlay und Füllung mit den Noten für die inhaltlich entsprechenden Teilprüfungen des praktischen Examens korreliert (siehe Tabelle

**Tabelle 1: Korrelation der Gesamtnoten und Teilnoten für Füllung und Inlay beider Prüfungsszenarien (\*\*. Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant.)**

	Examen Inlay	Examen Füllung	OSPE Füllung
Examen Füllung	.246** (r < .000)	/	/
OSPE Füllung	.042 (r = .538)	.123 (r = .067)	/
OSPE Inlay	.133* (r = .048)	.162* (r = .016)	.360** (r = .000)

1). Diese beiden Teilprüfungen wurden ausgewählt, da sie wegen großer inhaltlicher Überschneidungen einen relativ direkten Vergleich zulassen und ein Zusammenhang der Noten nahe legen. Hierbei ließ sich, unter Kontrolle der Kohorte als möglicher Einflussfaktor, feststellen, dass die Noten für Füllung in OSPE und Examen nicht signifikant miteinander korrelierten ( $r=.12$ ,  $p=.07$ ). Die Noten für Inlays zeigten dagegen einen solchen Zusammenhang ( $r=.13$ ,  $p=.05$ ).

Innerhalb einer Prüfung korrelierten die beiden Noten jeweils hoch miteinander. Das bedeutet, dass, wer gut (oder schlecht) in einem Teil der Prüfung abgeschnitten hat, auch im anderen Teil entsprechend gute (oder schlechte) Leistungen zeigte. Das spiegelt die gute Reliabilität der OSPE Gesamtprüfungen wieder (siehe Tabelle 1).

## 3.2. Genderunterschiede

### 3.2.1. Prüfungsnoten und Studiendauer

Bei 141 untersuchten weiblichen und 82 männlichen Studierenden konnte innerhalb der OSPE kein statistisch signifikanter Unterschied hinsichtlich des Geschlechts in den Teilnoten für Füllung und Inlay (indirekte Restauration) oder der Gesamtnote festgestellt werden. Auch bezüglich der Teilnoten und Gesamtnote im Staatsexamen zeigte sich kein Geschlechterunterschied. Ebenfalls bei der Studiendauer. Lediglich bei den Noten für Abitur und Physikum konnten Geschlechterunterschiede ermittelt werden. Im Bereich der Abiturnoten zeigte sich, dass weibliche Studierende ( $M=2.12$ ) signifikant besser Noten hatten ( $F=6.09$ ,  $p=.01$ ,  $\eta^2=.03$ ) als männliche Studierende ( $M=2.31$ ). Es konnten jedoch keine Korrelationen der Abiturnoten mit OSPE oder Staatsexamen gefunden werden, und somit lässt sich der Erfolg im OSPE beziehungsweise im praktischen Staatsexamen nicht durch die Abiturnoten vorhersagen. Dies kann in der Nutzung der Abiturnote als Selektionskennwert bei der Zulassung zum Studium begründet liegen, was deren Varianz in der Stichprobe deutlich einschränkt. Im Physikum (Zahnärztliche Vorprüfung) zeigte sich, dass männliche Studierende mit einem Mittelwert von 2.24 signifikant besser ( $F=4.15$ ,  $p=.04$ ,  $\eta^2=.02$ ) abgeschnitten hatten als weibliche Studierende ( $M=2.40$ ). Bei den männlichen Studierenden korrelierte diese Note zudem signifikant ( $r=.24$ ,  $p=.03$ ) mit dem OSPE, wobei eine leistungsstärkere OSPE-Note mit einer leistungsstärkeren Physikumsnote einherging, bei

den weiblichen nicht ( $r=.12$ ,  $p=.17$ ). Es konnten dagegen keine signifikanten Korrelationen zwischen Physikumsnote und Examensnote festgestellt werden.

### 3.2.2. Prädiktion der Staatsexamensnoten durch OSPE

Betrachtet man auch die Korrelation des OSPE getrennt nach Geschlecht der Studierenden, zeigt sich, dass bei den weiblichen Studierenden eine höhere Korrelation ( $r=.195$ ,  $p=.02$ ) zwischen OSPE und Note im Staatsexamen zu finden ist als bei den Männern ( $r=.137$ ,  $p=.22$ ). Ebenfalls zeigt sich eine hohe Korrelation ( $r=.27$ ,  $p=.007$ ) zwischen Studiendauer und Examensnote lediglich bei den weiblichen Studierenden, wenn nach Geschlechtern getrennt gerechnet wird. Die Berechnung getrennter multipler Regression mit schrittweisem Hinzufügen dieser beiden Variablen ergab, dass diese lediglich bei den weiblichen Studierenden einen signifikanten Beitrag zur Vorhersage der Examensnote leisten ( $\beta=.18$ ,  $p=.03$ ,  $R^2=.10$ ). Bei männlichen Studierenden war dieser Beitrag geringer ( $p=.16$ ,  $R^2=.05$ ).

## 4. Diskussion

Die im internationalen Kontext gut bekannte Prüfungsform des OSCE weist Reliabilitätswerte zwischen Cronbachs  $\alpha$  .40 und .91 auf [19]. Dies lässt sich sowohl für den rein medizinischen als auch für den dentalen Bereich feststellen [10], [18]. OSPE stellt eine mehr auf praktische Kompetenzen ausgerichtete Variation der OSCE dar [10]. Beide Begriffe werden in der Literatur häufig gleichbedeutend verwendet [10].

Studien zur prädiktiven Validität von OSCE/OSPE aus der Medizin und Zahnmedizin [8], [15], [17], [18], [20] belegen, dass die Leistungen der Studierenden in OSCE deren spätere klinische Kompetenz reflektieren. Unsere Studie bestätigt für die besondere Form einer OSPE eine hohe Reliabilität und eine prädiktive Validität für den Studienerfolg im praktischen Staatsexamen im Fach Zahnerhaltungskunde und steht somit im Einklang mit den oben zitierten Studien. Im Gegensatz dazu konnte in einer zahnmedizinischen Studie [21] keine signifikante Korrelation zwischen OSCE und Staatsexamen nachgewiesen werden. Eine mögliche Ursache wäre darin zu sehen, dass dort die Gesamtnote des Examens untersucht wurde (bestehend aus praktischem und mündlichem Examen), während unsere Studie den direkten Vergleich der prak-

tischen Fertigkeiten thematisierte. Es sind keine weiteren Publikationen bekannt, die OSCE/OSPE und Staatsexamen bezüglich Noten für identische Teilprüfungen untersucht haben. In diesem Zusammenhang stellt die vorliegende Studie eine Novität dar. Überraschend ist dabei, dass der Zusammenhang zwischen OSPE und praktischem Staatsexamen bei Frauen wesentlich deutlicher ist, was eine Reihe weiterer interessanter Fragen aufwirft. Dabei darf aber die Limitation der vorliegenden Studie nicht vergessen werden, dass die potentiellen Einflussfaktoren „Aufgaben“ und „Prüfer“ auf Basis der vorliegenden Daten, mit den verwendeten Methoden nicht mit analysiert werden konnten.

Des Weiteren ließ sich in unserer Arbeit feststellen, dass zwar die Noten für Inlays, jedoch nicht für Füllung bei OSPE und Examen signifikant miteinander korrelierten. Dies könnte an den unterschiedlichen Prüfungssettings (OSPE versus Staatsexamensprüfung) liegen. Es bleibt jedoch festzuhalten, ob nun die Prüfung unter klinischen Bedingungen an Phantompatienten (OSPE) oder an Patienten (Examen) stattfindet, die Aufgaben Füllung und Inlay offenbar ein sehr ähnliches Maß an Transferleistung und handwerklichen Fertigkeiten innerhalb einer Prüfung erfordern. Innerhalb einer Prüfung korrelierten diese beiden Teilnoten jeweils miteinander. Zwischen den Prüfungen scheinen sich die Anforderungen bei Inlay allerdings ähnlicher zu sein als bei der Füllung. Hier könnte vermutet werden, dass das räumliche Vorstellungsvermögen, dem generell bei Ausübung praktischer Tätigkeiten eine gewichtige Rolle zuteilwird, einen gemeinsamen Einfluss ausübt.

Die generelle Steigerung der praktischen Kompetenzen von der OSPE (im Mittel Note 3.79) hin zu der Staatsexamensbewertung (im Mittel 1.87) kann aus unserer Sicht auf einen erwünschten Trainingseffekt zurückgeführt werden. Diese Beobachtungen stimmen mit Ergebnissen der Studie von Sloan überein [20].

Weiterhin zeigte sich, dass weibliche Studierende signifikant bessere Abiturnoten erzielt haben, obwohl die vorliegende Spannweite der Noten durch ihre Nutzung als Kriterium für die Zulassung zum Studium stark eingeschränkt ist. Ähnliche Ergebnisse zeigt eine Analyse nach Auswertung der Daten von über 126 000 TeilnehmerInnen, die die Abiturprüfung bestanden und damit die formale Voraussetzung für ein Universitätsstudium erreicht hatten [<http://www.it.nrw.de> zitiert am 22.10.2015]. Festzuhalten bleibt jedoch, dass die Abiturnoten keine Korrelationen mit OSPE- oder dem praktischen Staatsexamensbewertungen aufwiesen.

Bei männlichen Studierenden war dagegen im Vergleich die Physikumsnote besser und es konnte eine Korrelation zwischen OSPE-Note und Physikumsnote festgestellt werden, die zeigt, dass eine leistungsstärkere OSPE-Note mit einer leistungsstärkeren Physikumsnote einhergeht. Dieser Zusammenhang ließ sich bei den weiblichen Studierenden nicht finden. Zu diesen Ergebnissen existieren keine publizierten und somit auch keine zitierbaren Vergleichsstudien. Tatsache ist, dass im Unterschied zur OSPE-Note, die die praktische Fertigkeit bewertet, es sich

bei der Physikumsnote um ein Konglomerat aus vier mündlichen Beurteilungen (Anatomie, Physiologie, Biochemie, Zahnersatzkunde) handelt. Die Note für Zahnersatzkunde umfasst auch eine praktische Bewertung, die sich aber im Rahmen unseres Studiensettings retrospektiv nicht herausfiltern ließ. Welcher gemeinsame Faktor oder Faktoren beide Noten bei männlichen Studierenden beeinflusst, kann daher nicht gesagt werden. Ferner ergaben die Zusammenhangsanalysen unserer Untersuchung eine Korrelation zwischen Studiendauer und Leistung im praktischen Examen. Eine längere Studiendauer korrelierte signifikant mit einer leistungsschwächeren Examensnote. Identische Ergebnisse wurden in einer zahnmedizinischen Untersuchung von Eberhard bestätigt [21]. Eine längere Studiendauer könnte dazu führen, dass die Motivation der Studierenden leidet bzw. die Einbindung in die Semesterzugehörigkeit derart gestört wird, dass von dieser ausgehende positive Auswirkungen verloren gehen. Zudem könnte die längere Studiendauer per se ein Hinweis auf externe Einflussfaktoren sein, die sowohl die Abschlussnote als auch die Dauer des Studiums beeinflussen. Denkbar wäre hier beispielsweise die finanzielle oder familiäre Situation der Studierenden. Das würde dazu passen, dass der Einfluss nur bei weiblichen Studierenden zu finden ist, wenn nach Geschlecht getrennt berechnet wird. Es konnte bereits gezeigt werden, dass diese stärker durch Mehrfachbelastungen wie Studium und Familie betroffen sind [22]. Bei den vorliegenden Ergebnissen zeigen sich insgesamt, durch die geringe und mittlere Höhe der Korrelationen sowie durch geringe Effektstärken, mehrere Hinweise darauf, dass weitere Einflussfaktoren im Modell fehlen. Das ist nicht weiter verwunderlich, da hier lediglich Leistungskennwerte und ähnliche Variablen betrachtet wurden. Der gesamte Bereich der Personenmerkmale fehlt bisher. Auch bei den Merkmalen der (Prüfungs)Situation sind die Daten nicht vollständig. Grund für beide Limitierungen ist die retrospektive Erfassung der Daten. Weiterhin als kritisch anzusehen ist die Tatsache, dass bei der vorliegenden Studie ausschließlich komplette Fälle erfasst wurden, was die Validität der Schlussfolgerungen aus den statistischen Analysen limitieren kann.

## 5. Fazit

Hinsichtlich der erzielten Ergebnisse kann man festhalten, dass die Prüfungsform der OSPE den Erfolg im praktischen Teil des Staatsexamens im Fach Zahnerhaltungskunde teils prädiziert, wenn auch die Ergebnisse erwartungsgemäß zusätzliche Einflussfaktoren wie die Studiendauer zeigen und weitere, aus den vorliegenden Daten nicht ersichtlichen, vermuten lassen. Für zukünftige Forschung wäre es daher wünschenswert, solche Faktoren das Studium begleitend zu erheben.

## Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

## Literatur

- Biggs J. Enhancing teaching through constructive alignment. *High Educ.* 1996;32(3):347-364. DOI: 10.1007/BF00138871
- Abou-Rass MA. Clinical evaluation instrument in endodontics. *J Dent Educ.* 1973;37(9):22-36.
- Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9 Suppl):S63-67. DOI: 10.1097/00001888-199009000-00045
- Thornton S. A literature review of the long case and its variants as a method of assessment. *Educ Med J.* 2012;4(1).
- Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med.* 2003;138(6):476-481. DOI: 10.7326/0003-4819-138-6-200303180-00012
- Hauer KE, Kohlwe J, Cornett P, Hollander H, Cate ten O, Ranji SR. Identifying entrustable professional activities in internal medicine training. *J Grad Med Educ.* 2013;5(1):54-59. DOI: 10.4300/JGME-D-12-00060.1
- Gadbury-Amyot CC, McCracken MS, Woldt JL, Brennan RL. Validity and reliability of portfolio assessment of student competence in two dental school populations: a four-year study. *J Dent Educ.* 2014;78(5):657-667.
- Graham R, Bitzer LAZ, Anderson OR. Reliability and Predictive Validity of a Comprehensive Preclinical OSCE in Dental Education. *J Dent Educ.* 2013;77(2):161-167.
- Ratzmann A, Wiesmann U, Kordaß B. Integration of an Objective Structured Clinical Examination (OSCE) into the dental preliminary exams. *GMS Z Med Ausbild.* 2012;29(1):Doc09. DOI: 10.3205/zma000779
- Schmitt L. Reliabilität der OSPE (Objective Structured Practical Examination) in Abhängigkeit von der Prüferanzahl und den Bewertungskriterien im ersten klinischen Semester des Faches Zahnerhaltungskunde, Promotion. Frankfurt: Goethe Universität Frankfurt am Main; 2014.
- Hallberg J. Quantitative Analyse der OSCEs eines Phantomkurses der Zahnerhaltung und Parodontologie, Promotion. Hannover: Medizinische Hochschule Hannover; 2012.
- Ortwein H, Fröhmel A, Burger W. Einsatz von Simulationspatienten als Lehr-, Lern und Prüfungsform. *Psychother Psych Med.* 2006;56:23-29. DOI: 10.1055/s-2005-867058
- Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *BMJ.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
- Nayak V, Baiyya KL, Adiga S, Shenoy S, Chogtu BM, Amberkar M, Kumari KM. OSPE in Pharmacology: Comparison with the conventional Method and Students' Perspective Towards OSPE. *Br Biomed Bull.* 2014;2(1):218-222.
- Martin IG, Jolly B. Predictive validity and estimated cut score of an objective structured clinical examination (OSCE) used as an assessment of clinical skills at the end of the first clinical year. *Med Educ.* 2002;36(5):418-425. DOI: 10.1046/j.1365-2923.2002.01207.x
- Smith LJ, Price DA, Houston IB. Objective structured clinical examination compared with other forms of student assessment. *Arch Dis Child.* 1984;59:1173-1176. DOI: 10.1136/adc.59.12.1173
- Schoonheim-Klein ME, Habets LL, Aartman IH, van der Vleuten CP, Hoogstraten 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(4):226-235. DOI: 10.1111/j.1600-0579.2006.00421.x
- Brown G, Manogue M, Martin M. The validity and reliability of an OSCE in dentistry. *Eur J Dent Educ.* 1999;3(3):117-125. DOI: 10.1111/j.1600-0579.1999.tb00077.x
- Cohen R, Reznick RK, Taylor BR, Provan J, Rothman A. Reliability and validity of the objective structured clinical examination in assessing surgical residents. *Am J Surg.* 1990;160(3):302-305. DOI: 10.1016/S0002-9610(06)80029-2
- Sloan DA, Donnelly MB, Schwartz RW, Felts JL, Blue AV, Strodel WE. The use of objective structured clinical examination (OSCE) for evaluation and instruction in graduate medical education. *J Surg Res.* 1996;63(1):225-230. DOI: 10.1006/jsre.1996.0252
- Eberhard L. Der erste Eindruck zählt? – Prädiktive Validität einer vorklinischen OSCE. Vortrag auf der AKWLZ-Tagung in Ulm. Ulm: Universität Ulm; 2014.
- Middendorf E, Apolinarski B, Poskowsky J, Kandulla M, Netz N. Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2012. 20. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch das HIS-Hochschul-Informationssystem. Bonn: Bundesministerium für Bildung und Forschung; 2013.

### Korrespondenzadresse:

PD Dr. med. dent. Susanne Gerhardt-Szép, MME  
Goethe-University, Dental School (Carolinum), Department  
of Operative Dentistry, D-60596 Frankfurt am Main,  
Germany, Phone: +49 (0)69/6301-7505, Fax: +49  
(0)69/6301-3841  
s.szep@em.uni-frankfurt.de

### Bitte zitieren als

Petko P, Knuth-Herzig K, Hoefler S, Stehle S, Scherer S, Steffen B, Scherzer S, Ochsendorf F, Horz H, Sader R, Gerhardt-Szép S. The reliability and predictive validity of a sixth-semester OSPE in conservative dentistry regarding performance on the state examination. *GMS J Med Educ.* 2017;34(1):Doc10. DOI: 10.3205/zma001087, URN: urn:nbn:de:0183-zma0010873

### Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2017-34/zma001087.shtml>

**Eingereicht:** 23.10.2015

**Überarbeitet:** 20.10.2016

**Angenommen:** 09.11.2016

**Veröffentlicht:** 15.02.2017

### Copyright

©2017 Petko et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.