The use of the objective structured clinical examination (OSCE) in dental education

Klein, M.E.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Implementing an Objective Structured Clinical Examination (OSCE) in dental education: effects on students’ learning strategies

M. E. Schoonheim-Klein¹, L. Habets², I. Aartman³, C. P. M. van der Vleuten⁴, J. Hoogstraten³ and U. van der Velden¹

¹Department of Periodontology, ²Department of Orthodontics, and ³Department of Social Dentistry and Behavioural Sciences, ACTA, the Netherlands.
⁴Department of Educational Development and Research, University of Maastricht Maastricht, the Netherlands

Abstract
Objectives: To investigate the effect of an Objective Structured Clinical Examination (OSCE) on dental students’ learning strategies and competence to manage periodontal diseases in patients. The implemented OSCE was expected to be superior to the existing Written Exam in fostering the acquisition of clinical competencies in terms of study strategies that are more oriented towards clinical practice, longer study time, greater clinical proficiency, and more realistic self-assessment.

Material and methods: After a clinical course in periodontology, 72 third year dental students were assessed summatively, either using a Written Exam or an OSCE (P-OSCE). The students were informed beforehand about the assessment formats. The self-assessed clinical competence, study time and strategies (i.e. practice with a manikin, peers and patient case) were evaluated by means of a questionnaire. After a comprehensive dental care course, all 72 students were assessed by an overall end-of-year OSCE, in which three periodontal stations were included ‘measuring pockets’, ‘educating patients’ and ‘tracing an X-ray with bone-loss’. The competence of the previous Written Exam group and the P-OSCE group was investigated by determining the mean scores and pass-fail scores of three periodontal test-stations as well as the total score of the end-of-year OSCE. The degree of realistic self-assessment was studied by correlating the self-assessed competencies as evaluated by means of the questionnaire with the total score of the end-of-year OSCE.

Results: Self-assessed clinical competence, study time and study strategies showed no differences between the P-OSCE and the Written Exam-group. The clinical competence determined in the test-station ‘measuring pockets’ in the end of year overall OSCE was higher for the P-OSCE group (P = 0.05) when compared with the Written Exam group; the two groups performed equally well in the test station ‘educating patients’, whereas the performance in ‘tracing an X-ray with bone-loss’ was better in the Written Exam group. This group also had a higher total score in the end-of-year OSCE (P = 0.05). The degree of realistic self-assessment was higher in the P-OSCE group than in the Written Exam group: in the P-OSCE group the self-assessed clinical competencies correlated significantly with the total score of the overall end-of-year OSCE (P ≤ 0.05). Conclusions: No effects of the implementation of an OSCE in undergraduate periodontal education were observed in study strategies, but the implementation of an OSCE in undergraduate periodontal education appears to stimulate learning, resulting in greater achievement of specific clinical competence and a greater level of realistic self-assessment.
Introduction

Studies have indicated that students’ conceptions of and approaches to learning are influenced by the method of assessment used (1–3). Students tailor their study activities to what is assessed rather than what teachers assert to be important. In other words, assessment drives learning (4). Although these are accepted principles in education, more empirical research is needed to clarify the educational effects of clinical assessment in order to determine how it can best be used to foster desirable learning habits (5, 6). Newble (7) found that changes in assessment methods in medical education affected the amount of time students put into training activities. Mavis (8) reported, however, that prior academic performance was a better predictor of the outcomes of an Objective Structured Clinical Examination (OSCE) than time spent on preparation. Within the framework of an overall reappraisal of strategies for assessing clinical competence in dentistry (9), a number of dental schools have introduced an OSCE (10, 11). As explained in a previous paper (12), the implementation of the OSCE was intended to serve several purposes: the OSCE was expected to improve alignment with the goals of clinical training and facilitate structured, objective assessment (9) and in addition the OSCE was expected to reveal students’ strengths and weaknesses, and thereby enhance the potential for learning (1, 5).

In Miller’s (13) competence pyramid, assessment moves from the ‘knows’ level via ‘knows how’ (Written Exams) and ‘shows how’ (performance in simulated conditions, such as OSCEs) to the final ‘does’ level of actual performance in daily practice. As Mavis (8) stated ‘Unlike the familiar paper-and-pencil tests, OSCEs have different performance expectations, and thus might require different studying strategies’ (p. 808). Gerrow et al. (14) compared the OSCE and Written component of the National Dental Examination Board of Canada. They showed that scores on written tests and OSCEs were positively correlated, but the scores on the Written Exam explained only 30% of the variance in OSCE scores. This level of association was expected by Gerrow et al. as they argued the two kinds of exams tap into different areas of competence.

Another way in which OSCEs might impact on learning was suggested by Shumway and Harden (15), who claimed that an OSCE provides ‘formative evaluation as the student is participating in it’. In other words, participation in an OSCE can enhance students’ learning.

Since written exams in a clinical context assess the ‘knows how’ level and OSCEs assess the ‘shows how’ level, such as communication and clinical skills, substituting an OSCE for a Written Exam might be expected to change the direction of students’ learning. Although it is generally assumed that assessment drives learning, this relationship has not been much studied in medical education, possibly because it is highly context dependent (16). Likewise, in dental education, the effect of a change in clinical assessment methods on students’ learning strategies has not yet been studied empirically.

Therefore, we investigated in this study the effects of an OSCE on students’ learning strategies and competence in treating patients with periodontitis by testing the hypothesis that, as compared to a traditional Written Exam, the OSCE would improve students’ learning of clinical competencies in treating patients with periodontitis by fostering more realistic self-assessment, more clinical-practice oriented study strategies, more time being spent studying, and better learning outcomes as regards measured clinical competence.
Materials and methods

Sample and context and the design of the study
A sample of 72 third-year dental students participated in a periodontal course. The students were randomized by the dental school administrative office to either a group that participated in a summative periodontal OSCE (P-OSCE), i.e. the experimental group (n = 36), or to a group that took part in a traditional summative Written Exam (n = 36), i.e. the control group. The mean age at the start of the academic year was 25.2 years (SD 4.8) in the P-OSCE group and 23.5 years (SD 2.6) in the Written Exam group. The percentage of women in the P-OSCE group and the Written Exam group were 42% and 58%, respectively. There were no significant differences between the P-OSCE group and Written Exam groups in age, gender and ethnicity.

The exams were scheduled at the end of the periodontal course. The course offered 1 day of practising clinical skills on a manikin and two seminars on diagnostics and the planning and initial treatment of periodontitis, including patient education. Thereafter, the students treated patients with periodontitis for a period of 10 weeks, supervised by staff of the Department of Periodontology.

Procedure and study design
At the start of the periodontal course, the students in the experimental group, the P-OSCE group, were informed about what the OSCE entailed and its procedure, the topics of the test-stations, and the assessment criteria. The students in the control group, the Written Exam group, were informed about the topics and the criteria of the Written Exam. This information was also available on the Internet in Blackboard.

Table 1 shows the study design: the periodontal course was taught in two periods of the year and students were assessed in March and in June. The course was taught on two different days of the week to two groups (A and B) that involved different teaching staff. Therefore, the students of the P-OSCE group and the Written Exam group were randomly allocated to either group A or B, in order to control for effects of any differences in teaching. All students who participated in this study also attended a course on comprehensive dental care. At the end of the third year, overall clinical competence in general dentistry was assessed by an overall end-of-year OSCE.

The P-OSCE was constructed in accordance with the guidelines developed by Brown et al. (9). Table 2 provides the blueprint of competence domain and station distribution across the P-OSCE. The selected domains of competence in this study were based on the profiles of the European dentist (17): I Knowledgebase, information handling and critical thinking, II Clinical information gathering and Diagnosis, III Treatment planning, IV Establishment and maintenance of oral health, with the major competence: managing periodontitis in patients, V Verbal communication and Health promotion, and VI Written communication. The test-stations were directly observed and rated by staff using checklists that required yes/no responses to 10 pre-determined items of adequate performance, yielding test scores per station ranging from 1 to 10. Five minutes were allocated for each station. Table 3 shows an example of the criteria and scoring of a test-station.
TABLE 1. Study design. Scheduling of third year periodontal and comprehensive care courses with assessments methods and dependent variables

<table>
<thead>
<tr>
<th>Period</th>
<th>Course</th>
<th>n</th>
<th>Assessment method</th>
<th>Measurement methods of dependent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>December-March</td>
<td>Perio. Day A</td>
<td>19</td>
<td>P-OSCE</td>
<td></td>
</tr>
<tr>
<td>December-March</td>
<td>Perio. Day B</td>
<td>15</td>
<td>Written Exam</td>
<td></td>
</tr>
<tr>
<td>March</td>
<td></td>
<td></td>
<td>Self-assessment</td>
<td>Questionnaire, VAS-scale and 6-point scale</td>
</tr>
<tr>
<td>April-June</td>
<td>Perio. Day A</td>
<td>21</td>
<td>Written Exam</td>
<td></td>
</tr>
<tr>
<td>April-June</td>
<td>Perio. Day B</td>
<td>17</td>
<td>P-OSCE</td>
<td></td>
</tr>
<tr>
<td>June</td>
<td></td>
<td></td>
<td>Self-assessment</td>
<td>Questionnaire, VAS scale and 6-point scale</td>
</tr>
<tr>
<td>September-June</td>
<td>Comprehensive</td>
<td>105</td>
<td>Overall OSCE</td>
<td>Perio stations scores Overall End-of-year OSCE total score</td>
</tr>
</tbody>
</table>

TABLE 2. Full blueprint of the P-OSCE according to station number and competence domain tested and test-station contents and sequence used in P-OSCE

<table>
<thead>
<tr>
<th>Competence domains</th>
<th>Test stations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Knowledge</td>
<td>v</td>
</tr>
<tr>
<td>Information gathering and Diagnostics</td>
<td>v</td>
</tr>
<tr>
<td>Treatment planning</td>
<td>v</td>
</tr>
<tr>
<td>Establishing health</td>
<td>v</td>
</tr>
<tr>
<td>Verbal Communication and health promotion</td>
<td>v</td>
</tr>
<tr>
<td>Written communication</td>
<td>v</td>
</tr>
</tbody>
</table>

1, history taking of simulated patient with periodontitis; 2, tracing an X-ray picture with bone loss; 3, measuring pockets on manikin; 4, interpretation of periodontal chart: prognosis and diagnosis; 5, educating simulated patient about the use of inter-dental brushes; 6, rest; 7, scaling and root planing of molar on manikin; 8, writing a referral letter; 9, writing a prescription for antibiotics; 10, interpretation of periodontal chart in combination with X-rays; 11, interpretation of re-evaluation of the periodontal therapy; 12, rest; 13, put in order a protocol for initial periodontal therapy.

The Written Exam was based on a documented case of a periodontitis patient (periodontal charts, slides and X-rays), which was posted on Blackboard 1 week before the examination. The examination consisted of five essay questions about diagnosis and formulation of a treatment plan. In addition, the students were asked to trace the image of bone level and teeth on a dental X-ray. The written exam took 1 h.

Furthermore, immediately after the P-OSCE or the Written Exam, the students completed a questionnaire to evaluate the impact of the different exams on students’ self-assessment of clinical competencies, on additional study time, and study strategies. The questionnaire was not anonymous, because the responses had to be linked to the examination results.
Two days after the examinations the students were informed about their performance in the examinations. Furthermore, they received global feedback of their competence in managing periodontal diseases in patients. The global feedback consisted of a 10 min meeting with each of the students and included a discussion on their strengths and weaknesses based on their OSCE performance.

**Evaluation of self-assessed competence**

Self-assessment of competence was evaluated in two ways. First, the students were asked to indicate their competence levels for six competence domains on a 4-point scale: 1, assistant; 2, novice; 3, intermediate; and 4, competent. The six domains were the same as those tested in the P-OSCE, i.e. knowledge (four questions), history taking and diagnosis (seven questions), treatment plan (three questions), establishing health (six questions), verbal communication (four questions), and written communication (three questions). The scores were summed per domain. Second, the students were asked to rate their overall competence on a visual analogue scale (VAS) ranging from 0 to 100 mm in response to the question: ‘Do you feel competent after this course to carry out initial periodontal therapy in periodontitis patients?’

**Evaluation of additional study effort and study time and study strategies**

The amount of time students spent studying for the examination was investigated with two questions. Students were asked to rate on a VAS (0–100 mm) whether they had put additional effort into preparation for the examination. They were asked ‘Did you put in additional effort to prepare for the exam beyond the scheduled course time in the clinic?’ Second, they were asked to indicate their response to the question: ‘How much additional time did you spend on preparation for this exam beyond the scheduled course time in the clinic?’ on a 6-point scale: 1, 0–30 min; 2, 31–60 min; 3, 61–120 min; 4, 121–240 min; 5, 241–480 min; and 6, >480 min. These time scores can be grouped in three equivalent groups: 1 and 2, 1 h or less; 3 and 4, 2–4 h; 5 and 6, >4 h.

Study strategies were evaluated by asking the students how they had prepared for diagnostic competence, clinical skills and communication skills, respectively. They could choose from eight strategies: practising with fellow students, practising with others than students, practising on a
manikin, practising with patient cases, revision using ACTA manuals, textbooks, literature/articles and questioning staff members. For each skill, one or two of the methods were deemed to represent appropriate clinical-practice oriented study strategies. For diagnostic competence, this was practising with patient cases, for clinical skills, it was practising on a manikin and for communication skills it was practising with fellow students or others.

Measurement of proficiency of clinical competence in the overall end-of-year OSCE
The overall end-of-year OSCE consisted of 17 test stations and three rest stations. The 17 stations pertained to cariology, endodontontology, periodontology, prosthetics, radiology, orthodontics, surgery and infection control. The test stations were directly observed and rated by staff using criterion-based checklists that required yes/no responses to 10 predetermined items of adequate performance, yielding test scores per station ranging from 1 to 10. Five minutes were allocated for each station.

The P-OSCE group and the Written Exam group were compared on the total score of the overall end-of-year OSCE and on the only three periodontal stations that were included in the overall end-of-year OSCE, i.e. 1, measuring pockets; 2, patient education; 3, tracing an X-ray with bone-loss. The pass-fail score for each station was set at 5. The total score of the overall end-of-year OSCE could range from 0 to 170.

Degree of realistic self-assessment
Finally, the P-OSCE group and the Written Exam group were compared with respect to their degree of realistic self-assessment. As a measure of the degree of realistic self-assessment, we correlated the self-assessed competence in the six domains with the total score of the overall end-of-year OSCE.

Statistical analysis
For statistical analysis SPSS11.0 was used. ANOVA was performed to examine differences between the P-OSCE group and the Written Exam group in age, summed scores of self-assessed competence per domain, self-assessed overall competence (VAS), additional study time (VAS), and total scores on the overall end-of-year OSCE by controlling for effects of gender, ethnicity (native Dutch, vs. non-native Dutch), and periodontal course group A or B and periodontal examination-month March or June. Mann-Whitney U-tests were used to examine differences between the P-OSCE and the Written Exam group for gender and ethnicity, and for study time (6-point scale). Chi-square tests were performed to test for effects in gender and ethnicity. Also, chi-square tests were used to test differences between the two examination groups in study strategies and pass–fail scores in the overall end-of-year OSCE and in addition, logistic regression analysis was used to test for interactions with gender, ethnicity, and group A or B in the periodontal course.

To examine whether self-assessments per domain were realistic, Spearman rank correlation coefficients were calculated for self-assessed competence per domain and total score on the end-of-year OSCE for the two groups. P-values ≤ 0.05 were considered as statistically significant.
Results

Self-assessed competence
Table 4 shows no significant differences between the P-OSCE group and the Written Exam group in self-assessed competencies in the six domains. However, evaluation of the self-assessed overall competence in treating periodontal patients (VAS, 0–100 mm) showed that students in the P-OSCE group scored lower (mean = 81.3, SD = 9.4) than the students in the Written Exam group (mean = 85.8, SD = 9.9; P = 0.04).

Additional study effort and study time and study strategies
The analysis showed that students in the P-OSCE group and the Written Exam group reported they put equal amounts of additional study effort into preparation for their respective assessments (Table 5). In comparison with native Dutch students, non-native Dutch students put more additional study effort into preparation for the P-OSCE than for the Written Exam. With regard to the additional study time, the results showed that students in the P-OSCE group and the Written Exam did not differ in self-reported minutes of additional study time beyond the scheduled clinical course time (Table 6). Most students studied for between 61 and 120 min, with the two examination groups using about the same amount of time. In contrast with native Dutch students, non-native Dutch

Table 4. Realistic self-assessment: Spearman correlation coefficients (and p-values) and mean scores (SD) between self-assessed competencies and the total score of the Overall End-of-year OSCE in the P-OSCE group and the Written Exam group

<table>
<thead>
<tr>
<th>Domains of competence</th>
<th>P-OSCE group</th>
<th>Written Exam group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-assessment</td>
<td>Total-score end-of-year OSCE</td>
</tr>
<tr>
<td>Knowledge</td>
<td>11.9 (1.8)</td>
<td>112.4 (17.7)</td>
</tr>
<tr>
<td></td>
<td>n=34</td>
<td>n=34</td>
</tr>
<tr>
<td>Information gathering and diagnosis</td>
<td>22.5 (2.3)</td>
<td>112.8 (17.9)</td>
</tr>
<tr>
<td></td>
<td>n=34</td>
<td>n=34</td>
</tr>
<tr>
<td>Treatment planning</td>
<td>9.2 (1.2)</td>
<td>112.1 (18.1)</td>
</tr>
<tr>
<td></td>
<td>n=36</td>
<td>n=36</td>
</tr>
<tr>
<td>Establishing health</td>
<td>16.2 (1.9)</td>
<td>111.8 (18.2)</td>
</tr>
<tr>
<td></td>
<td>n=35</td>
<td>n=35</td>
</tr>
<tr>
<td>Verbal communication and health promotion</td>
<td>14.0 (1.7)</td>
<td>114.0 (17.1)</td>
</tr>
<tr>
<td></td>
<td>n=31</td>
<td>n=31</td>
</tr>
<tr>
<td>Written communication</td>
<td>4.1 (1.5)</td>
<td>113.8 (16.8)</td>
</tr>
<tr>
<td></td>
<td>n=31</td>
<td>n=31</td>
</tr>
</tbody>
</table>

*p≤ 0.05 level
**p≤ 0.01 level
students spent more additional time preparing for the P-OSCE than for the Written Exam. Table 7 shows students' use of appropriate study strategies for each competence. Statistical analysis showed that students in the P-OSCE group did not report more clinical-practice oriented strategies than did the students in the Written Exam group.

**Proficiency of measured clinical competence in the overall end-of-year OSCE**

The results of three periodontal test stations in the end-of-year OSCE are shown in two ways: mean scores and the pass–fail scores (Table 8-I and II). With regard to the mean scores of the three periodontal tests–stations, the Written Exam group scored higher on 'tracing an X-ray' compared with the P-OSCE group. There was no difference in mean scores for the other two periodontal test-stations in the end-of-year OSCE. However, there was a statistically significant interaction effect for periodontal course group A or B: the students of the Written Exam group in group A had better results on 'tracing an X-ray' than the P-OSCE group students in group A; the students of the P-OSCE group and the Written Exam group in group B demonstrated no differences in performance on this competence.

There was no statistical interaction between examination groups and the month in which the exams were taken (March or June) in mean scores on the three periodontal test-stations or in the total score in the end-of-year OSCE.

With regard to the pass–fail scores (Table 8-II), more students in the P-OSCE group passed the overall end-of-year OSCE test station on 'measuring pockets' and the end-of-year OSCE test station on 'tracing an X-ray' was passed by more students in the Written Exam group. There were no statistically significant differences between the two groups for the test-station on 'patient education'. Analyses showed that no interaction effects were present for gender, ethnicity, or periodontal-examination-month with regard to the pass–fail scores.

Finally, the Written Exam group had a higher end-of-year OSCE total-score than the P-OSCE group (Table 8-III). Further analysis showed that group B students from the Written Exam group obtained higher end-of-year OSCE total scores than the group B students from the P-OSCE group.

**The correlation between the self-assessed competencies and the total score in the end-of-year OSCE as a measure of the degree of realistic self-assessment**

The total scores on the end-of-year OSCE showed significant positive correlations with all self-assessed competencies of each of the six competence domains in the P-OSCE group only (Table
4). There was no effect of ethnicity or periodontal course group. However, there was a gender effect. In the P-OSCE group female students showed a significant positive correlation between self-assessed clinical skills and the total score of the end-of-year OSCE, whereas males showed a significant positive correlation between the self-assessed communication, treatment planning and knowledge, and the total score of the end-of-year OSCE. In the Written Exam group, the female students showed a significant positive correlation between self-assessed clinical skills and verbal communication with the end-of-year OSCE total score, whereas the male students showed no significant correlations of end-of-year OSCE total-score with any of the self-assessed competencies.
**Discussion**

After an OSCE, we expected the students to be more prudent and realistic with regard to self-assessed competence. This assumption was confirmed significant correlations between self-assessed competencies and the end-of-year OSCE scores were found for the P-OSCE group, but not for the Written Exam group. Participation in the preceding OSCE gave students the opportunity to perform in different clinical domains and gain experience of their clinical competence. This was true for all items except for the tracing of bone on a radiograph, which also tests performance. The Written Exam tested further exclusively the ‘knows how’ level in Miller’s Pyramid (13) of knowledge and diagnostic skills, reasserting Gerrow’s view (14) that these two examinations tap into different areas.

The expectation that students would prepare for the P-OSCE with more practice-oriented study strategies was not confirmed: there were no immediately apparent effects of the implementation.
Chapter 3

of an OSCE manifested in students’ study strategies. In the P-OSCE and the Written Exam group about 50% used a case report to study diagnostics. Surprisingly, 20 students in the Written Exam group said that they used a manikin to practise skills, just as many as in the P-OSCE group (Table 7). An explanation could be that they have misinterpreted this question. The two groups used the same strategy for studying communication, even though the students knew, or could have known, that communication would not be included in the Written Exam. The finding that the groups did not differ in study strategies is in agreement with previous results (8). In addition, Mavis’ observation (8) that the amount of study time was highly variable was also confirmed in our study. Since this was the first OSCE in the periodontal department, there were no experienced peers to inform the students about what to expect. Apparently, the lecture on the methods of the OSCE and the information on the Internet were not sufficient to direct students’ study strategies and study time. The implicit intervention of directed study strategies might not have been sufficiently strong to change existing study strategies. However, this is likely to change in the near future as students become more familiar with the format and the results and the procedure of the OSCE are becoming better known in the dental school and more incorporated in the curriculum.

That assessment can direct students’ learning was shown by the differences in learning outcomes between the P-OSCE and the Written Exam group in the overall end-of-year OSCE. The P-OSCE group performed better on clinical competence, i.e. measuring pockets, than did the Written Exam group. The practice in the P-OSCE itself after the periodontal course may have helped the students to perform better in the next overall end-of-year OSCE. However, although a previous OSCE can be a predictor for the results on the same clinical skills in the next OSCE, this effect cannot be generalised to other skills. Thus, although the experimental OSCE students were better in measuring pockets than the control students in the Written Exam group, the P-OSCE group performed less well than the Written Exam group on tracing Xrays. With regard to this test station, student performance showed an effect of the different periodontal course groups, i.e. an effect related to different teaching staff. Teachers do indeed continue to play a prominent role in respect of students’ learning. This study shows that teachers have an impact on the learning process even in a student-centred learning environment. Additional conversations with teaching staff confirmed this finding.

The finding that the students of the Written Exam group in group B performed better on the end-of-year OSCE than the students of the P-OSCE group in group B is not easily explained. However, it could be that the clinical competencies learned in the periodontal course are only a predictor for the periodontal competencies in the end-of-year OSCE, and not for competencies in other clinical areas like Cariology and Endodontology. As van der Vleuten stated, selecting an assessment method involves context-dependent compromises: the predominant condition affecting the reliability of assessment is domain- or content-specificity, because competence is highly dependent on context or content (16, 18). From the results of our study, we conclude that the content-specificity problem might also be true in the effect or impact of assessment on learning, also termed as ‘educational validity’ (19, 20). Further research is needed to confirm this theory.
There was no immediate feedback given directly after each individual test-station during the P-OSCE, but global feedback was given after 2 days. Immediate feedback can be highly influential on students’ learning and performance, as Hodder (21) claimed, it can improve the students’ competency in the performance of criterion-based tasks. Still, it is possible that the feedback given after 2 days had a positive influence on the P-OSCE group students and improved their competency in the overall end-of-year OSCE. Additionally, individual experiences of the students on the periodontal course combined with the existing student study strategies might account for some of the results. Also, the differences in self-assessment of the P-OSCE group and the Written Exam group might partly be attributable to their changes in perception brought about by their personal practical experience and feedback received on their performance during the 10 week course.

The effect of the P-OSCE on realistic self-assessed competence appeared to be different for males and females. Studies of gender differences in self evaluation show ambiguous results. Ericson et al. (22) found Swedish female students to score slightly more effectively, whilst underrating occurred more frequently than over-rating. Arnold et al. (23) found no association in the USA between self-ratings and gender, but also found that students at each level, except year 3, rated themselves lower than did their docents. Also, in the USA, surgical residents demonstrated greater underestimation of clinical skills amongst female residents than amongst their male colleagues (24). In another study in Sweden (25), the self-assessment abilities of dental students in a periodontal course were studied in the domains knowledge, diagnostics, communication and clinical skills. They found no gender differences in self-assessment of competence in these domains, although females scored better in the examinations than males. It might be possible that in our study as well as in the Swedish and USA studies, female students underestimated themselves.

In the present study, although this was not explicitly investigated, it was found that gender and ethnicity were biases in the demonstration of clinical competence. The female students outperformed their male counterparts on the end-of-year OSCE. This confirmed one of Wiskins’ et al. (26) and Haq et al. (27) findings, suggesting that medical female students in the UK in an OSCE are better performing under examination conditions. Furthermore, the nonnative Dutch students in the P-OSCE and the Written Exam group differed from the native Dutch students in the amount of study time and extra study effort devoted. It seems logical that language problems will require non-native Dutch students to spend more time on preparation than native Dutch students. Additional explorative analyses with ethnicity showed that non-native Dutch students scored statistically significantly lower on the total score in the end-of-year OSCE; this is in line with the findings of Friedman et al. (28). Further research is needed to explore the causes of underperforming of male foreign students in OSCEs.

**Conclusion**

The hypothesis that the implementation of an OSCE would have an effect on students’ learning was supported by the finding of more realistic self-assessment as well as more proficiency of specific clinical competence in managing periodontitis in patients. In terms of study strategies
no immediate effects of the implementation of an OSCE in undergraduate periodontal education were observed. However, from this study, we learned that there is a different influence of the OSCE and the Written Exam on students' learning outcomes. OSCE stimulated learning in the dental clinic and gave the students a greater level of realistic self-assessment. These results support the well accepted statement that assessment drives learning.

Acknowledgements
The authors wish to thank Marc Timmermans for the fruitful discussions and Mereke Gorsira for her language suggestions.
References

24. Minter RM, Gruppen LD, Napolitano KS, Gauger PG. Gender differences in the self-assessment of