The use of the objective structured clinical examination (OSCE) in dental education
Schoonheim-Klein, M.E.

Citation for published version (APA):
Chapter 4

Language background and OSCE performance: a study of potential bias

M. Schoonheim-Klein¹, J. Hoogstraten¹, L. Habets¹, I. Aartman¹, C. van der Vleuten², M. Manogue², U. van der Velden².
¹ ACTA, Amsterdam, the Netherlands
² University of Maastricht, the Netherlands
³ Leeds Dental Institute, Leeds, UK

European Journal of Dental Education 2007; 11: 222–229
Chapter 4

Abstract

Aim: To investigate whether Non-Native dental students perceive drawbacks in their education and examination experience because of their lack of language proficiency and to test whether prolonging Objective Structured Clinical Examination (OSCE) test-station duration improves their performance.

Materials and methods: Dental students (n = 345) completed a questionnaire about their native country, their language background and the possible drawback they perceived in dental education and examination because of their language proficiency. Students were marked as 'native', when they were born in the Netherlands with Dutch as native language or 'Non-Native' when they were born outside the Netherlands, raised with a non-Dutch native language, or raised bilingually. A sample of 108 students was assessed by an OSCE testing a periodontal course with nine test-stations. Test-station topics were: (1) history taking, (2) measuring attachment level, (3) educating patients, (4) tracing a radiograph, (5) root-planing, (6) writing a prescription, (7) diagnostics and prognostics, (8) differential diagnostics and (9) writing a referral letter. The first five test-stations mentioned were of 5-min duration. The other four test-stations were provided in two modes: either with a short (5 min) or longer (10 min) version. Every student took at random two long and two short test-stations.

Results: In the group of 345 questionnaire responders, Non-Native students (n = 116) perceived significantly more drawback in education and examination than Native students (n = 229) (P < 0.001). When Non-Native students speak Dutch at home, around 38% of them reported perceived drawbacks in education, whereas when they speak their native language at home, around 60% reported perceived drawbacks in education (P = 0.005). In the periodontal OSCE (n = 108), the Native group (n = 70) had significantly higher total scores than the Non-Native (n = 38), (P = 0.009, d = 0.53). The Non-Native group had significantly lower mean scores in the communication station ‘educating patients’ (P = 0.034, d = 0.42). Prolonged test-station duration from 5 to 10 min had no positive effect in all experimental test-stations in the Native and Non-Native group. Female students in the Native group out-performed male in a communication test-stations. Female students in Native and Non-Native groups were found to be more successful in ‘tracing bone loss on radiographs’.

Conclusion: Non-Native students perceived a drawback in dental education and examination because of their language proficiency in Dutch, which is confirmed by their actual OSCE performance. Prolonging the time for a test-station did not improve OSCE performance of Non-Native students. It is recommended that students with problems in language ability need additional tuition and practice.
Introduction

Effective communication is a key issue in both education and health care. One of the factors of communication is language proficiency. For example, limited language proficiency can be a barrier in gaining academic success. It has been shown that, the lower the level of foreign language proficiency the more important a factor it becomes in determining academic success in foreign students (1, 2). Furthermore, the quality of communication between patients and clinicians can have a major impact on the quality of care and treatment outcomes (3, 4).

Communication skills in the healthcare professions are often tested in an Objective Structured Clinical Examination (OSCE). In this mode of examination, the examinee circulates through a series of ‘stations’, which test isolated elements of clinical competence. In medical education in the UK, studies have shown poorer performance for ethnic minorities in OSCEs (5), in particular in communication stations (6). Additionally, in the USA (7), a high correlation between the clinical competence assessment and the spoken English proficiency was found. In contrast, Bienstock et al. (8) did not find the effects of ethnicity on OSCE results. However, in this retrospective study, they were unable to establish fluency in the English language. Yet, Australian studies (9, 10) showed that students, who preferred using a language other than English at home, performed at a lower level in oral communication examinations than those students who preferred to use English. Lumb and Vail (11) concluded that the relatively poor performance of male and ethnic minority students in medical OSCEs needed urgent further investigation. From other studies, it appears that when cases are too difficult for student’s language ability, this will contribute to construct-irrelevant variance in the assessment. Thus, ratings in performance-based assessments, like OSCEs, may be influenced by potentially unfair biases as gender, ethnicity and social status. Such construct irrelevant sources of variation should be prevented. The OSCE should be fair to everyone and every student should have appropriate training to achieve the same standard of competence (12–14). The OSCE should not measure language background, but measure the competences.

In dentistry, OSCEs are also used for testing diagnostic, clinical and communication skills, all based on relevant knowledge (15–17). In a previous study in our school (18), the performance of Non-Native dental students proved to be significantly lower than the performance of Native Dutch dental students. This could be due to lower proficiency in the Dutch language. It could be postulated that the level of language proficiency of Non-Native Dutch students entering the dental school should, ideally, be sufficient to start an academic study, as they are tested by the Dutch language proficiency test: ‘Test of Dutch as a second language.’ This test is compulsory for students with a foreign secondary education to test Dutch language proficiency prior to entering a Dutch university (19). The test assesses for the level B2 of Common European Framework of Reference for Languages. The test is comparable with the TOEFL (Test of English as a Foreign Language) used to test English language proficiency of foreign students in the UK, USA and Canada. Still it might be possible that the initial level of language proficiency decreases, when Non-Native students continue to speak their native language at home, or that a test result belies the actual proficiency in practice. In our previous study (18), as well as in other studies in medical education (5, 20), women tend to perform better in clinical assessments, also within
Chambers (21) found that examinees with a non-English native language spent significantly more time with the patient than those whose native language is English. Therefore, it might also be possible that Non-Native Dutch students need more time in an OSCE test-station. In dental OSCEs, the test-station-length is often set at 5 min (16, 17), yet there is rarely any supporting rationale provided for the choice of this time limit. Van der Vleuten and Swanson (22) could not discern an obvious relationship between test-station length and reliability. They recommended that the length of a test-station should be determined by the test-station task. It has been shown however, that increasing test-station length from 5 to 10 min, decreases the generalisability in the 10-min test-stations (23). Advantages of a short testing time would imply lower costs of the examination or allow more test-stations in the allotted examination time and a potential improvement in the reliability.

The purpose of this study is to investigate whether Non-Native dental students perceive drawbacks in education and examination because of their lack in language proficiency; and to test whether longer OSCE stations affect performance.

Materials and methods

Language proficiency and perceived drawback
Clinical undergraduate dental students (n = 430) in the years 2005 and 2006 were asked to complete a questionnaire (see Fig. 1) about their native country, their language background and the possible drawback they perceived in dental education and examinations because of their language proficiency. The questions were related to their language background, their native language, being raised bilingually and the language used at home. The perceived drawback in dental education was assessed using a Likert scale, with a scale from 1, ‘no drawback’ to 5, ‘very much drawback’. Drawback is defined by the perceived negative effect or restriction by language ability. The lack of language proficiency was not tested but instead the student’s own perception of the drawback resulting from his lack of language ability was measured.

Students were categorised as Native, when they were born in the Netherlands with Dutch as native language. Students were categorised as Non-Native, when they were born outside the Netherlands, or raised with a non-Dutch native language, or raised bilingually. The questionnaire was tested on intrarater reliability (Kappa = 0.87) and is shown in Fig. 1.

The questionnaires were coded so that results could be linked to later performance in dental examinations. The students gave informed verbal consent and were guaranteed that the outcome would not affect their progress through the dental programme. Completed questionnaires were received from 345 students.

Student performance and perceived drawback
A sample of 108 third year students out of the 345 questionnaire respondents were assigned by the administration of the dental school to participate in a periodontal course in 2005 and 2006. The content of the courses and the assessments in both years were identical and given by the same staff.
The periodontal course was described in an earlier paper (17) and was developed to teach the competencies to perform an initial treatment of a periodontal patient. At the start of the periodontal course, the students were informed about what the OSCE entailed, the topics of the test-stations, and the assessment criteria. The course offered 1 day to practise 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. At the completion of the course, the students were assessed by a summative OSCE with short and long stations.

The periodontal OSCE

The OSCE of the periodontal course was constructed in accordance with the guidelines developed by Brown et al. (13) and was designed with a blueprint as described in an earlier paper (17). The OSCE had nine test-stations and one rest-station. We used trained actors performing as simulated patients. Test-stations’ topics were (see Table 1): 1. history taking of a simulated patient, 2. measuring attachment level on a manikin, 3. educating a standardised patient to use inter-dental brushes, 4. tracing bone-loss on a radiograph, 5. scaling and root-planing on a manikin, 6. writing a prescription, 7. diagnostics and prognostics, 8. differential diagnostics and 9. writing a referral letter. The first five test-stations mentioned were each of 5-min duration. The other four test-stations were experimental, with both a short (5 min) and a long (10 min) version. The students

Please answer the following questions concerning your language background:

Question 1. Is Dutch your primary language?

Question 2. Were you being raised bi-lingual?
   If yes, which languages did you speak at home as a child?

Question 3. Which language do you normally speak during dinner at home?

Question 4. Which language did you speak yesterday during dinner?

Question 5. (Please circle the number most appropriate)
Do you feel restricted by your language understanding and language ability in dental education?

1  2  3  4  5
No drawback        Very much drawback

Question 6. (Please circle the number most appropriate)
Do you think you are negatively affected because of your language ability, when taking exams?

1  2  3  4  5
No drawback        Very much drawback

Question 7. Do you have dyslexia?

Question 8. Do you have a certificate of dyslexia?

Remarks:

Fig. 1. Questionnaire on language background.
Chapter 4

... did not take the same station twice in a long and short version but each student was randomly assigned to two long and two short test-stations (route A or B). Figure 2 shows an example of 'by-passes' in the route of the OSCE designed for this experiment. The test-stations were rated by staff using checklists that required yes/no responses to 10 pre-determined items of adequate performance, yielding test scores per test-station ranging from 1 to 10.

### Statistical analysis

For statistical analysis, SPSS11.0 was used (SPSS Inc., Chicago, IL, USA). Mann–Whitney U-tests were used to compare differences between the two language groups, Native and Non-Native, in terms of any perceived drawback in dental education and examination (Likert scale). Wilcoxon signed ranks tests were used to test for differences between drawback in education and examination in the two language groups (total group and men and women separately). A Mann–Whitney U-test was also used to test for differences in perceived drawback within the Non-Native group between

---

**TABLE 1. Full blueprint of the OSCE according to test-station number and competence domain tested**

<table>
<thead>
<tr>
<th>Competence domains</th>
<th>Test-stations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>9</td>
</tr>
<tr>
<td>History taking and Diagnostics</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment planning</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Clinical skills</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Verbal communication</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Written communication</td>
<td></td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>


---

Fig. 2. An example of the design of one of the two by-passes in the OSCE for the experiment with Short (5 min) and Long (10 min) test-stations.
students who spoke Dutch at home and students who spoke in their native language at home. ANOVA was performed to examine differences between the Native and Non-Native group in station performance and total scores on the OSCE. Interaction effects with gender were also considered using a two-way ANOVA. In the Native and Non-Native group, Spearman rank correlations were performed to correlate the level of perceived drawback and the OSCE performance. T-tests were used to test for differences of the test-station score between long and short OSCE test-stations in the Native group and the Non-Native group separately. P < 0.05 was accepted as statistically significant. Finally, the effect size was calculated using the method described by Cohen (24). Effect size is used to quantify the degree of practical significance (25).

Effect size d = (mean 1-mean 2)/SD, pooled. According to Cohen, effect sizes of 0–20, 0.50 and 0.80 are considered as a moderate, a medium or crucial practical importance respectively.

Results

Language proficiency and perceived drawback

The group of 345 dental students included 229 Native and 116 Non-Native students, 37% and 42% male students respectively. The most common language spoken at home by the Non-Natives was Dutch (56%) followed by 21% languages from the middle east, 8% Asian languages (e.g. Indonesian, Vietnamese and Chinese), 6% Slavic languages (Yugoslavian and Russian), and some European languages. Both Native and Non-Native students reported a perceived drawback (score >1) in education (7% and 47% respectively) and in examinations (8% and 47% respectively) because of their language proficiency. In both groups, a small percentage of the students reported severe drawbacks (2% and 10% respectively, as shown in Fig. 3.). Analysis of the data showed a significant difference between the Native and Non-Native students in perceived drawback in education and examination (P < 0.001). The results for men and women, when treated separately, yielded the same result.

Non-Native students felt a greater drawback in performing in examinations than in dental education (P = 0.02), whereas for Natives there was no difference.

When Non-Native students spoke Dutch at home, around 38% of them reported perceived drawbacks (score >1) in education, whereas when they spoke their native language at home, around 60% reported perceived drawbacks in education (P = 0.005).

Student performance, language proficiency and perceived drawback

The 108 students participating in the periodontal course included 30% males in the Native group and 42% males in the Non-Native group. The majority (66%) of Non-Native students were speaking their native language at home and reported perceived drawbacks in education (63%) and examinations (63%) because of language proficiency. The Native students perceived drawback in education (11%) and in examinations (16%) because of language proficiency.

The results of the evaluation of the periodontal course by means of the OSCE are shown in Table 2. The total score of the Native students was significantly higher than that of the Non-Native
Fig. 3. Percentage of Native and Non-Native dental students perceiving drawbacks in (a) education and (b) examination.
group (P = 0.009) and the effect size ‘d’ was found to be 0.53. In eight out of nine test-stations there was no significant difference in OSCE performance between the Native and Non-Native students. Only in the test-station ‘educating a patient in the use of interdentally brushes’, did the Non-Native students score significantly lower than the Native students (P = 0.034, d = 0.42). For the test-stations, ‘diagnostics and prognostics’, ‘differential diagnosis’, and ‘writing a referral letter’, a trend was observed towards a better performance from the Native students, but this failed to reach the level of significance. In addition, there appeared to be an interaction between gender and ‘history taking’. In the Native group, female students outperformed male students in the communication test-station, i.e. history taking from a simulated patient (P = 0.015). In both the Native and Non-Native group, female students outperformed male students in tracing bone-loss on a radiograph (P = 0.041).

In the Non-Native group, in the group of students speaking their native language at home, the OSCE performance in the communication station ‘educating patients’ and the level of perceived drawback in education as well as in examination was statistically significant and negatively correlated (r = -0.429 P = 0.025 and r = -0.409 P = 0.034 respectively). With none of the other OSCE test-stations statistically significant correlations were found.

In the Native group, only performance in the test-station ‘diagnostics and prognostics’ and the level of perceived drawback in education and examination were statistically significant (positively) correlated (r = 0.306 P = 0.010 and r = 0.296 P = 0.013).

<table>
<thead>
<tr>
<th>Test-station</th>
<th>Native students N=70</th>
<th>Non-Native students N=38</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. History taking</td>
<td>7.39 (1.5)</td>
<td>6.87 (1.7)</td>
<td>0.098</td>
</tr>
<tr>
<td>2. Measuring attachment level</td>
<td>6.30 (2.0)</td>
<td>6.08 (1.8)</td>
<td>0.574</td>
</tr>
<tr>
<td>3. Educating patients</td>
<td>8.00 (1.2)</td>
<td>7.45 (1.4)</td>
<td>0.034*</td>
</tr>
<tr>
<td>4. Tracing bone loss on radiograph</td>
<td>5.47 (1.2)</td>
<td>5.61 (1.7)</td>
<td>0.645</td>
</tr>
<tr>
<td>5. Scaling and rootplaning</td>
<td>6.94 (1.8)</td>
<td>6.63 (2.1)</td>
<td>0.427</td>
</tr>
<tr>
<td>6. Writing a prescription †</td>
<td>9.17 (1.4)</td>
<td>9.34 (1.0)</td>
<td>0.590</td>
</tr>
<tr>
<td>7. Diagnostics and prognostics †</td>
<td>7.51 (1.7)</td>
<td>6.95 (1.4)</td>
<td>0.079</td>
</tr>
<tr>
<td>8. Differential diagnosis †</td>
<td>6.47 (2.3)</td>
<td>5.50 (2.8)</td>
<td>0.055</td>
</tr>
<tr>
<td>9. Writing referral letter †</td>
<td>6.34 (2.0)</td>
<td>5.58 (2.0)</td>
<td>0.066</td>
</tr>
<tr>
<td>Total score</td>
<td>63.60 (6.7)</td>
<td>60.00 (6.8)</td>
<td>0.009**</td>
</tr>
</tbody>
</table>

†, short (5 min) and long (10 min) station
*p < 0.05
**p < 0.01
Chapter 4

The effect of prolonging the duration a test-station from 5 to 10 min to improve students’ performance was investigated in four test-stations, i.e. ‘writing a prescription’, ‘diagnostics and prognostics’, ‘differential diagnosis’ and ‘writing a referral letter’. For both Native (data not shown) and the Non-Native students (see Table 3), no statistically significant effect was found.

Discussion
In our previous study (17) it was found that students born outside the Netherlands performed less well in an OSCE than Native Dutch students. It was suggested that language proficiency might be partly responsible for this effect. The findings of the present study confirmed this supposition: the Non-Native students perceived more drawbacks in education and examination because of their language proficiency. Also, there was an association between the level of perceived drawback in Non-Native students who spoke their native language at home, with performance in the test-station: ‘educating patients’. In addition, the Non-Native students performed less well, both in the OSCE total score and in those stations that required them to demonstrate the education of patients. The effect size of 0.53 as found in this study reaches the level of medium practical importance as defined by Cohen (24). Therefore, the drawback Non-Native students perceived because of their language proficiency resulting in the underperformance of the OSCE should be a serious concern for the dental school. This may be illustrated by the remark of the simulated patient that some Non-Native students were really very good in communication whereas others were incomprehensible.

The Non-Native and the Native students were both very cooperative in giving the information about their language abilities. The response rate was 80%. Both groups made the comment that language proficiency is a serious problem in the dental school.

The few Native Dutch students that perceived drawbacks because of language abilities probably suffer from learning disability. Two Native students with serious drawbacks (Fig. 3) mentioned a certificate relating to dyslexia. It is plausible that dyslectic students will profit from prolonged test-stations, but the number of dyslectic students was too small to analyse this in the present study.

The length of a test-station should be determined by the test-station task (22). In the present study, the tasks in the test-stations were developed to fit in 5 min. And although students often complain about the lack of time in a 5-min test-station, we did not find a difference in scores in

---

**TABLE 3. Mean values (standard deviation) of the performance of Non-Native students (n=38) in the long and short test-stations of the periodontal OSCE.**

<table>
<thead>
<tr>
<th>Test-station</th>
<th>5 min</th>
<th>10 min</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Writing a prescription</td>
<td>9.33 (0.9)</td>
<td>9.35 (1.0)</td>
<td>0.959</td>
</tr>
<tr>
<td>7. Diagnostics and prognostics</td>
<td>6.85 (1.0)</td>
<td>7.06 (1.7)</td>
<td>0.656</td>
</tr>
<tr>
<td>8. Differential diagnosis</td>
<td>5.19 (2.6)</td>
<td>5.73 (3.0)</td>
<td>0.566</td>
</tr>
<tr>
<td>9. Writing referral letter</td>
<td>5.73 (2.3)</td>
<td>5.38 (1.9)</td>
<td>0.614</td>
</tr>
</tbody>
</table>
the 5- or 10-min test-stations for Native and Non-Native Dutch students. For example, in the
test-station where students had to write a referral letter, this was difficult for Non-Native Dutch
students (a mean score of 5.58 out of 10 was achieved), but their performance did not improve in
the 10-min version. Therefore, we assume that, if language ability is a problem, simply giving more
time is not the solution to better phrasing.

In the literature (5, 20) and in our previous study (17), it was suggested that female students
in an OSCE perform at a higher level than males. This was confirmed in the present study. In
the test-station 'communication', in the Native group, female students outperformed their male
counterparts. Furthermore, female students were better in the test-station: 'tracing bone-loss on
a radiograph' in both Native and Non-Native students. However, this gender effect could not be
found in the results of the OSCE overall.

The group of Non-Native students who spoke their native language at home perceived more
drawback than the Non-Native students who spoke Dutch at home. The present OSCE included
a small group of those Non-Native students (n = 11) speaking Dutch at home. This small group
(n = 11) mostly originated from Surinam, where Dutch is spoken at school, and they were raised
bilingually. A student's remark is that the Dutch language they were speaking at home is somewhat
different from the Dutch spoken in the Netherlands and this sometimes led to misunderstanding
and drawback in education and communication. The group was too small to compare their
performance with the Native students.

There is no doubt that the diversity in a society should be reflected in the population of dental
school students. Students experience this as an enrichment, as reported in many studies about
diversity in medical and dental education (26, 27). Non-Native students often do well academically
(28), but they perceive different problems than Native students. It would be wrong to assume
problems and differences in communication skills on the basis of an individual's ethnicity and
language proficiency alone. Also, factors as study style, personality, domestic environment, former
education and culture can have influence on success (5, 6). For these reasons, the results of this
study should be interpreted with some caution.

The language test 'Test of Dutch as a second language' is compulsory for students with a foreign
secondary education to test Dutch language proficiency prior to entering a Dutch university
(19). The level required in this test is 'B2'. When Non-Native students are speaking their 'mother
language' at home, this level may actually reduce in time, whilst they should reach a higher level
'C' = to speak, understand and study at an academic level. Therefore, it is advised that the Non-
Native students should take language lessons to improve their language abilities. Maybe the
implementation of a 'language buddy', a native Dutch speaker, could help to improve the level of
language proficiency.

**Conclusion**

Within the limits of the present study (e.g. variability amongst individuals in age, gender, cultural
background and former education) we conclude that Non-Native students perceive there to be
a drawback in their experience of dental education and dental examination because of language
proficiency, which is confirmed by in this example of actual performance in an OSCE. We conclude that the differences we observed were educationally significant and should be addressed. Prolonging the time for an OSCE test-station does not improve OSCE performance of Non-Native students and this is not a solution for the language problems. Rather, students with problems in language ability need additional tuition and practice in Dutch. In addition, they could be encouraged to speak Dutch away from the dental school during their dental education.
References

21. Chambers K, Boulet JR, Gary NE. The management of patient encounter time in a high-stakes assess-