Is there a difference between self-perceived performance and observed performance in an Objective Structured Clinical Examination (OSCE)? An exploratory study among medical students in the United Arab Emirates [version 1]

Erik Koornneef¹, Tom Loney², Ahmed R. Alsuwaidi³, Marilia Silva Paulo⁴

¹Erasmus School of Health Policy and Management
²College of Medicine
³Department of Pediatrics
⁴Institute of Public Health

Abstract
This article was migrated. The article was not marked as recommended.

Competency-based education and training has become a key component of healthcare systems across the globe. Ensuring that healthcare professionals are able to assess their own competencies is critical for continued professional development and the delivery of high-quality care. The aim of this study was to assess how medical students perceive their performance on an objective structured clinical examination. Using a cross-sectional study design, a sample of Emirati third and fourth year (preclinical) medical students (N=106; 56.4% response rate) was recruited from the United Arab Emirates University in Al Ain, United Arab Emirates. Medical students completed a short non-invasive clinical task (i.e. measuring and recording blood pressure and performing hand hygiene) followed by a structured survey to self-assess their performance and skills. Trained assessors used a clinical skills observation checklist tool to score each student’s performance. According to the observed performance, 27.36% of medical students performed the objective structured clinical task adequately. In contrast, 69.52% rated their own performance as adequate. Furthermore, only 8.43% of medical students rated their own clinical skills as below average. This study did not find evidence that medical students can accurately assess their own clinical skills
and performance. In order to support the delivery of high-quality healthcare, it is important that medical students develop their ability to accurately assess their own clinical skills and performance early in their medical careers. Teaching and appraising self-reflection is an important component of any undergraduate or postgraduate medical degree program.

Keywords
medical education, objective structured clinical examination, self-assessment, United Arab Emirates

Corresponding author: Marilia Silva Paulo (mariliap@uaeu.ac.ae)
Competing interests: No competing interests were disclosed.
Grant information: The author(s) declared that no grants were involved in supporting this work.
Copyright: © 2018 Koornneef E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
How to cite this article: Koornneef E, Loney T, Alsuwaidi AR and Paulo MS. Is there a difference between self-perceived performance and observed performance in an Objective Structured Clinical Examination (OSCE)? An exploratory study among medical students in the United Arab Emirates [version 1] MedEdPublish 2018, 7:180 https://doi.org/10.15694/mep.2018.0000180.1
Is there a difference between self-perceived performance and observed performance in an Objective Structured Clinical Examination (OSCE)? An exploratory study among medical students in the United Arab Emirates

Erik Koornneef[1], Tom Loney[2], Ahmed R. Alsuwaidi[3], Marilia Silva Paulo[4]

Abstract

Competency-based education and training has become a key component of healthcare systems across the globe. Ensuring that healthcare professionals are able to assess their own competencies is critical for continued professional development and the delivery of high-quality care. The aim of this study was to assess how medical students perceive their performance on an objective structured clinical examination. Using a cross-sectional study design, a sample of Emirati third and fourth year (preclinical) medical students (N=106; 56.4% response rate) was recruited from the United Arab Emirates University in Al Ain, United Arab Emirates. Medical students completed a short non-invasive clinical task (i.e. measuring and recording blood pressure and performing hand hygiene) followed by a structured survey to self-assess their performance and skills. Trained assessors used a clinical skills observation checklist tool to score each student’s performance. According to the observed performance, 27.36% of medical students performed the objective structured clinical task adequately. In contrast, 69.52% rated their own performance as adequate. Furthermore, only 8.43% of medical students rated their own clinical skills as below average. This study did not find evidence that medical students can accurately assess their own clinical skills and performance. In order to support the delivery of high-quality healthcare, it is important that medical students develop their ability to accurately assess their own clinical skills and performance early in their medical careers. Teaching and appraising self-reflection is an important component of any undergraduate or postgraduate medical education.
degree program.

Keywords: medical education; objective structured clinical examination; self-assessment; United Arab Emirates

Introduction

Medical education plays an important role in maintaining and improving the quality of a country's healthcare system (Khalid, 2008). Many competencies are defined for medical students that must be acquired before graduation, such as clinical knowledge and expertise, professional integrity, empathy, communicative skills, and conceptual thinking (Patterson et al., 2000; Huenges et al., 2017). To achieve these desired competencies, future doctors need to be able to accurately self-assess and appraise their multiple skills, also in addition to recognizing their limitations (Huenges et al., 2017). In this paper, we assumed that competency involves multiple skills. Healthcare providers and educators are moving towards competency-based education and assessment skills, and the lack of self-assessment skills from healthcare professionals can act as a barrier for self-paced learning (Graves, Lalla and Young, 2017). Self-assessment has multiple definitions in the literature and the term has also been used to describe self-reflection or self-evaluation. Andrade and Du (5, page 160) define each of these concepts independently, and in this paper we used their self-assessment definition as the "process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work and revise accordingly" (Andrade and Du, 2007). Studies have found that physicians often assess themselves as being more competent than they actually are (Davis et al., 2006). Therefore, introducing self-assessment for medical students may assist them to accurately assess their own skills and competencies in the future. Accurate self-assessment of personal and professional capabilities are now seen as essential for success (Alwi and Sidhu, 2013) as healthcare professionals and essential for delivery of high quality care.

The Objective Structured Clinical Examination (OSCE) is a comprehensive evaluation tool that has been used to assess the competencies of medical students in the majority of medical schools worldwide (Zayyan, 2011). The OSCE assesses clinical skills, counselling, and communication-based competencies through direct observation (Zayyan, 2011). The OSCE has been widely used over the past two decades and can be defined as a "timed examination in which medical students interact with a series of simulated patients in stations" (Zayyan, 2011). The OSCE comprises several clinical stations, usually 10-12, where the student performs tasks including history-taking, physical examinations, counselling or patient management, and clinical procedures. The student is required to complete the task within a set time limit and according to well-defined criteria for each specific clinical skill. These clinical tasks are normally assessed by trained assessors from the medical faculty (Zayyan, 2011; Kim, 2016).

This study took place in the United Arab Emirates (UAE), an independent federation, consisting of seven Emirates with a total population of approximately 9.1 million people, in 2016 (Federal Competitiveness and Statistics Authority, 2017). It is a relatively young, high-income country, established in 1971 (Abdel-Razig and Alameri, 2013) with a strong government-led desire to build a world-class healthcare system to improve the health of its population (Koornneef, Robben and Blair, 2017). The World Health Organization described the Eastern Mediterranean Region, where the UAE is located, as a region facing major challenges regarding the healthcare workforce. Specifically, the UAE faces major challenges related to the shortage of UAE national healthcare workers, a high reliance on expatriate staff, limited health professionals’ production capacity, and a high turnover of expatriate healthcare workers (World Health Organization, 2017). In this context, the present study focuses on one of these challenges: the capacity deficit to educate and train an adequate number of appropriately educated and trained UAE nationals’ healthcare professionals.
The main objective of the study was to explore the differences between self-assessment and trained-assessor OSCE score. Our hypothesis was that a medical student who rates their clinical skills and competencies as adequate would also achieve a higher observed OSCE overall score.

Methods

Study design

A cross-sectional study was used to investigate the relationship between self-perceived performance from medical students and trained-assessor rated OSCE performance. The STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) Statement was used to structure this paper (Elm et al., 2007).

Setting

The study was conducted at the clinical skills simulation centre of the College of Medicine and Health Sciences of the United Arab Emirates University, the largest public university in the UAE. Data collection occurred over two consecutive days in April 2016. This study was approved by the institution's Social Sciences Research Ethics Committee (ERS_2015_3212).

Participants

Medical students from the Doctor of Medicine (M.D.) six-year program at the College of Medicine and Health Sciences were the study population. Pre-clinical students (third and fourth year) were invited to participate in the survey and to perform a specific non-invasive clinical task (measure blood pressure).

Variables

The study variables were overall OSCE score, student self-assessed performance, and self-reported clinical skills. These two last variables were measured by statements in the survey, ranked by a Likert scale ranging from one to five. The variable self-assessed performance was defined through the survey sentence "Overall, I think that I performed the OSCE to the best of my abilities" measured by the Likert scale as strongly disagree (1), disagree (2), neither (3), agree (4) and strongly agree (5). The variable self-reported clinical skills was defined through the sentence "I would rate my own clinical skills and competence as" categorized into (1) poor, (2) fair, (3) average, (4) good and (5) very good.

The dependent variable overall OSCE score was created by summing the scores of the clinical skills observation tool that was completed by the observers. The trained observers were faculty and staff from the College of Medicine. They were considered eligible to assess the clinical task of collecting blood pressure by their qualifications, and they were professionally trained on how to evaluate the quality of hand hygiene practice, having successfully completed a two hour long online hand hygiene course from Hand Hygiene Australia and through a bespoke two-hour face-to-face practical course prepared by the authors.

Data sources/measurements

To accomplish our research objective, we used a cross-sectional survey and a clinical observation tool to collect the data. The survey was designed specifically for this study and the designing process took into consideration a review of other papers and surveys (Makkai and Braithwaite, 1996; Sunshine and Tyler, 2003; Murphy, Tyler and Curtis, 2009; Tyler, Mentovich and Satyavada, 2014). The survey formed part of a larger study exploring medical student's perceptions of healthcare regulation (Koornneef et al., no date) and included questions regarding the two above
mentioned variables (self-perceived performance and self-reported clinical skills).

The clinical skills observation tool was designed in consideration of other observation tools used to assess OSCE, for example, the OSCEstop (OSCEstop, 2013). This observation tool included data collection on four major parts: preparation, including introducing self to the patient, hand hygiene including the WHO hand hygiene standards (before and after the clinical task), and blood pressure measurement (clinical task performed at OSCE). These four parts were assessed by observers using a Likert scale ranging from one to three (one – performed adequately, two – attempted, but performed inadequately and three – not attempted).

Eligible medical students received an email invitation to participate in the research study one week before the study took place. Students were informed and asked to perform a clinical task and to complete the survey. Students who were willing to participate booked a slot or ‘walked in’ at the clinical skills simulation centre during the two days of the data collection. Upon arrival, the students received a brief description of the study and consent process, and they were requested to read and sign a consent form. A research assistant explained the study as follows: the participant was asked to perform a short non-invasive clinical task – measuring and recording a person’s blood pressure – and complete the survey afterwards. Students were randomly assigned to one of the four available clinical skills simulation rooms. One of the observers played the role of the "standardized patient", and the other one pretended to be completing a Sudoku book, but observed the student performing the OSCE and completed the clinical skills observation tool. Usually the OSCE is a circuit of stations, but as this OSCE was designed specifically for this study, it comprised only one station with one clinical task. At the end of the task, the participant was asked to complete the survey and earned a Certificate of Attendance. All students had received the same training on performing the clinical task and were aware of the key steps involved in completing the task correctly and in accordance with the UAE health regulations.

Bias

To minimize potential bias in our study, the observers were not known to the students, they were always of the same gender as the participants and they were trained and experienced in observing students’ OSCE performance. In addition, each participant was randomly assigned to the clinical room where the OSCE was carried out. The layout of the clinical observation rooms was identical. Students were unaware (blinded) to the covert assessor role of the research assistant who pretended to complete the Sudoku book whilst they performed their clinical task. This method of blinding was used to minimize any possible Hawthorne effect (i.e. observer effect that causes reactivity in which an individual modifies their behaviour in response to awareness of being observed).

Study sample size

All undergraduate medical students from the third and fourth year (N=188) were invited to participate in the study. Of the 188 students, 106 participated in our study (56.38% response rate).

Quantitative variables/Statistical methods

Descriptive statistical techniques were used to describe the dependent variable (trained assessor rated overall OSCE score) and the two independent ones under analysis: self-perceived performance and self-reported clinical skills. A t-test was used to test the difference between genders and the dependent variable. An ANOVA was used to determine the difference between the categories of the independent variables and the OSCE overall score. All the tests were performed using α=5%.
Results/Analysis

Participants

A total of 106 medical students participated in our study representing 31.80% of all undergraduate medical students in the university. All university students from the College of Medicine and Health Sciences at the United Arab Emirates University are UAE nationals, and 77.40% were female. The proportion of male/female in the study sample is similar to the gender distribution of the medical student population in the college.

Main results

When asked if they performed the OSCE to the best of their abilities, the majority (69.52%) of students answered agree or strongly agree, while nearly a third (30.48%) of students self-assessed their performance as neutral (neither) or negative (disagree or strongly disagree) (Figure 1).

Figure 1: Medical students self-perceived performance after OSCE.

Half of the students (55.66%) self-reported their clinical skills as ‘good’ and only 8.49% considered their clinical skills below average (Figure 2). None of the students rated their clinical skills and competencies as ‘poor’.
Figure 2: Medical students self-reported clinical skills.

The observed score shows that the OSCE overall score was performed ‘adequately’ by 27.36% of students, while 72.64% were rated as ‘attempted, but performed inadequately’. None of the students did not attempted.

The mean (±SD) of the trained-assessor observed OSCE overall score was 1.7±0.0, minimum of 1.0 and maximum of 2.6. The mean (±SD) of the trained-assessor observed OSCE score for females was 1.7±0.0 and for males was 1.6±0.0 (Figure 3). This difference was not statistically significant ($p=0.794$).
Figure 3: Overall OSCE score per gender.

The students that 'strongly disagreed' and the students that 'neither agreed nor disagreed' to performing the OSCE at their best had a mean OSCE overall score of 1.6±0.1 and 1.6±0.0, respectively (Figure 4). The students that 'strongly agreed', 'agreed' and 'disagreed' revealed same mean OSCE overall score with a decimal difference amongst them. ANOVA was calculated to assess the difference between the students’ perceived performance categories and there were no statistically significant differences (p=0.763).
Figure 4: Overall OSCE score and medical students self-perceived performance.

The students that reported their clinical skills as ‘fair’ showed the highest mean (±SD) OSCE overall score (1.8±0.1). While the students who reported their clinical skills to be ‘good’ or ‘very good’ presented mean (±SD) overall OSCE score of 1.7±0.0 and 1.7±0.1, respectively. There was no statistical significance between how students reported their clinical skills and OSCE overall score ($p=0.6$).
Figure 5: OSCE overall score and medical students self-reported clinical skills.

The intragroup variance between gender, self-perceived performance, and self-reported clinical skills is not statistically significant ($p=0.492$).

**Discussion**

**Key results**

The key result is that this study did not find evidence to support the hypothesis that medical students in the preclinical phase can accurately self-assess their own skills, competencies and performance. In other words, the lack of a statistical significance between the mean of overall OSCE scores the two self-rated variables may indicate that medical students in the preclinical phase have not yet developed the necessary self-reflection skills to accurately appraise their own performance compared to their assessed performance. There was no difference between the gender of the medical students regarding self-assessment and trained-assessor observed overall OSCE score. These findings were similar to Andrade and Du’s study that explored the attitudes toward and beliefs about self-assessment in undergraduate teacher education students in the United States and did not find differences in the responses of male and female students (Andrade and Du, 2007).

**Limitations**

The undergraduate preclinical medical students that participated in the present study represented nearly a third (31.80%) of the total medical students at the United Arab Emirates University. One of the limitations of this study is that it represents a convenient sample from one of six medical universities in the UAE, and includes only third and
fourth-year preclinical medical students.

**Interpretation**

Only one-quarter of preclinical medical students performed the OSCE adequately. However, the majority of the students reported a positive self-assessment when asked if they performed the OSCE to their best ability. In Oman, a similar study compared the difference between the student’s self-assessment and the trained-assessor OSCE score in 60 medical students and the results show that the students consistently overestimated their performance in four of the 12 items while underestimating their performance in the remaining eight items (Jahan *et al.*., 2014).

Almost 70% of participants self-reported their clinical skills as good or very good and that they had completed the OSCE to the best of their ability. This is in stark contrast with the actual trained-assessed OSCE appraisal which found that only 27% of students performed the OSCE task adequately. Other studies have found similar discrepancies. In a systematic review including 20 studies on the accuracy of physician self-assessment compared with observed assessments, the results showed that physicians did not accurately self-assess themselves in the majority of the studies (Davis *et al.*, 2006). In addition, the systematic review reported only weak or no associations were found between self-rated assessment and external observed assessments (Davis *et al.*, 2006). The inaccuracy of self-assessment is also reported in medical students as being frequent and across several specialities or levels in the graduating program (Eftekhar *et al.*, 2012; Graves, Lalla and Young, 2017; Huenges *et al.*, 2017). The timing of assessment has been shown to play a role in student self-reflection. A study examining the self-rated competencies of 168 medical students pre- and post-OSCE showed that students decreased their self-rating after the family medicine objective examination, but not significantly for family medicine specific skills (Graves, Lalla and Young, 2017). A study of 244 medical students for the specialization in general practice revealed that the method of self-assessment was experienced and perceived as useful, but only 57% of the sample opted for self-assessment combined with individual feedback on their strengths and weaknesses (Huenges *et al.*, 2017). Self-assessment is a complex process of internalization and self-regulation (Andrade and Du, 2007), and many medical students may not have developed the necessary cognitive skills and reflective practices during their medical undergraduate degrees to provide a realistic self-appraisal. Therefore, providing sufficient time for students to develop their self-reflection skills is an important component of any undergraduate or postgraduate medical degree programme.

Some authors have questioned the reliability of self-assessment (Davis *et al.*, 2006; Chen *et al.*, 2008; Graves, Lalla and Young, 2017). It has been reported by medical students that if the subjective self-rating is to be used as a formal aspect of the medical education program, then it should be complemented with formative feedback from the supervisors (Huenges *et al.*, 2017). As such, several researchers advise the development of all-inclusive continuing professional education programs including portfolios, documenting practice-based learning and improvement activities, and creating less general and more detailed learning objectives (Davis *et al.*, 2006; Huenges *et al.*, 2017).

In this case, it is important to include direct observation in clinical training which has also been a standard in medical education as it is linked to students self-confidence in their final year (Chen *et al.*, 2008). For future studies including medical students, we would suggest including a third way of measuring clinical competencies: peer review, this would ensure a triangulated measurement: self, peer and external assessments (Colthart *et al.*, 2008).

**Conclusion**

The self-assessment of medical students is not related to trained-assessed OSCE score in this study. To achieve good practices in future healthcare professionals, specifically physicians, it is important to understand the discrepancies between the medical student’s self-perception and their actual observed performance. Further research is required to provide a deeper understanding of the factors related to the discrepancy between student self-assessment and trained-assessed performance. Such detailed information would allow educators to create better learning
environments with more effective self-assessment strategies. This paper contributes to the understanding of the current production of Emirati medical students in the UAE, to achieve the UAE Vision 2021 and to the 2030 agenda of the Sustainable Development Goals and Universal Health Coverage.

**Take Home Messages**

Only a quarter of all students performed the OSCE task adequately, but almost 70% of participants self-reported their clinical skills as good. Study findings did not provide evidence to support the hypothesis that medical students can accurately self-assess their own competencies; however, study findings indicate that teaching self-reflection is an important component of any undergraduate or postgraduate medical degree programme.

**Notes On Contributors**

Erik Koornneef is a PhD candidate with the Erasmus School of Health Policy and Management, Erasmus University, Rotterdam, the Netherlands and the Head of Healthcare for Cognit, a Mubadala and IBM Watson joint venture. His research interest is focused on the effectiveness of regulation in healthcare and innovative ways to improve healthcare quality.

Dr Tom Loney is an Associate Professor in Public Health and Epidemiology in the College of Medicine at the Mohammed Bin Rashid University of Medicine and Health Sciences in Dubai, United Arab Emirates. Part of his research portfolio focusses on improving undergraduate and postgraduate medical, dental, and public health education.

Dr Ahmed R. Alsuwaidi is an Associate Professor of Pediatrics and Infectious Diseases in the College of Medicine and Health Sciences at the United Arab Emirates University in Al Ain, United Arab Emirates. His current research interests include vaccine-preventable diseases, tuberculosis, respiratory viral infections, and medical education.

Marilia Silva Paulo is completing her PhD in Health Policies and Development at the Institute of Hygiene and Tropical Medicine, NOVA University, Lisbon, Portugal. Her current research focusses on the development of healthcare services to improve chronic care services for patients with diabetes, cardiovascular diseases, and cancer.

**Acknowledgements**

None.

**Bibliography/References**


in Higher Education. Routledge, 32(2), pp. 159–181. [https://doi.org/10.1080/02602930600801928]


Appendices

None.

Declarations

The author has declared that there are no conflicts of interest.

This has been published under Creative Commons “CC BY 4.0” (https://creativecommons.org/licenses/by-sa/4.0/)

Ethics Statement

This study was approved by the institution’s Social Sciences Research Ethics Committee (ERS_2015_3212).

External Funding

This paper has not had any External Funding
Open Peer Review

Migrated Content

Version 1

Reviewer Report 02 June 2019

https://doi.org/10.21956/mep.19608.r29136

© 2019 Masters K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ken Masters
Sultan Qaboos University

This review has been migrated. The reviewer awarded 2 stars out of 5

An interesting paper discussing the differences between self-perceived performance and observed performance in an OSCE among medical students in the UAE. The authors begin by establishing the medical necessity of having medical students and professionals who are able to assess their own level of competency. The study design has been well constructed and is well-described. The authors should, however, have reflected more deeply on the implications of the differences and the possible causes. While the difference in self-reporting and assessing is certainly problematic, the validity of the assumption that the students have made the error is based upon the premises that (a) the assessors’ assessment is accurate, and (b) the assessors’ assessment matches what the students have been taught. (This is because the students' self-assessment will naturally be based upon a comparison of what they have been taught and what they did.) The troubling aspect to this, and one which calls for further exploration, is not the large difference: the main problem is that the assessors found only 27% of the students to be performing the task adequately. This would indicate that there is something seriously wrong. The error may be (a) teaching, or (b) a disjuncture between what the assessors are expecting and what the students have been taught, or (c) Others, which may include ill-suited students or incompetent assessors. Until this has been investigated, one could just as easily argue that the students are correct in their self-assessment, but the assessors are wrong. The only thing we know is a difference, but we have no way of knowing the cause of that difference. For one to conclude that the students are at fault, a more detailed study of the other alternatives is required. (Although the authors touch on this in their conclusion, the tone of the paper implies that, whatever the causes, they are on the students’ side). Some smaller issues with the paper: • It would have been preferable if the abstract had been laid out with the structured headings. • There is a line “None of the students did not attempted.” which needs addressing. So, a useful study, but the authors should have gone quite a bit deeper into investigating the possible causes of the discrepancy; at the very least, they should acknowledge that their study shows only
the discrepancy, and that one cannot conclude the reasons in any direction, and further investigation is necessary.

**Competing Interests:** No conflicts of interest were disclosed.

**Reviewer Report 05 December 2018**

https://doi.org/10.21956/mep.19608.r29137

© 2018 Hays R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Richard Hays
James Cook University

This review has been migrated. The reviewer awarded 3 stars out of 5

Thank you for this interesting paper which adds to the evidence base indicating that self-assessment in the higher stakes context of a formal assessment probably cannot be relied on. The finding is not new, but is documented in another context, suggesting that this phenomenon crosses cultures, education and health systems and nationalities. Just as most people self-rate their driving skills as superior, candidates in examinations often have a higher innate sense of ability that expert judges. Self-assessment may be more reliable in formative assessment, but in this case my interpretation is that the assessment was voluntary and not part of formal progress assessment - is that correct? Perhaps the authors could clarify if the observation took place after standardised training in the skill that was observed. This is not clear, yet the low rate of success on rater assessment (30%) raises questions about the teaching. One interesting issue that is not explored here, but could be, is the relationship between a gap between self and rater assessments and acceptance of feedback. That would earn more stars.

**Competing Interests:** No conflicts of interest were disclosed.