

OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE) AS A RELIABLE EVALUATION STRATEGY: EVIDENCE FROM A BRAZILIAN MEDICAL SCHOOL

Luciana Castellani

Bahiana School of Medicine and Public Health (EBMSP), Brazil
Salvador University (UNIFACS), Brazil
E-mail: lucianacastellani@gmail.com

Luiz Fernando Quintanilha

FTC University Center (UniFTC), Brazil
Salvador University (UNIFACS), Brazil
E-mail: quintanilha.lf@gmail.com

Maria Belen Arriaga

Oswaldo Cruz Foundation (FIOCRUZ), Brazil
E-mail: mbag711@gmail.com

Maria de Lourdes Lima

Bahiana School of Medicine and Public Health (EBMSP), Brazil
Salvador University (UNIFACS), Brazil
E-mail: mlourdeslima@bahiana.edu.br

Bruno Bezerril Andrade

Bahiana School of Medicine and Public Health (EBMSP), Brazil
Oswaldo Cruz Foundation (FIOCRUZ), Brazil
E-mail: bruno.andrade@fiocruz.br

Abstract

The new curriculum of medical schools has brought the need to develop a professional committed to ethics, capable of understanding the role of social factors, the insertion of the individual into the family and their role in promoting health. These skills and the use of new methodologies in teaching-learning process have demonstrated that the traditional performance assessment methods are limited in holistically evaluating the student. This study analyzed the correlation among different methods for evaluating competencies in medical internship students and Objective Structured Clinical Examination (OSCE). A cross-sectional, descriptive and analytical study was performed among 5th year medical students at Pediatric, Surgery, Medical Clinic, Family and Community Medicine and Gynecology and Obstetrics. Grades of each student obtained through different evaluations (theoretical test, conceptual evaluation and test skills) at the end of each rotation were correlated with the individual results from OSCE. In regard to correlations with the OSCE score values, it was found a weak to moderate positive and statistically significant relationship with the global cognitive score ($r = .22, p < .001$), global abilities score ($r = .26, p < .001$) and with the global behavioral score ($r = .38, p < .001$). The students' performance in the various assessments of performance evaluated here was positively correlated to their performance on OSCE. Furthermore, the organization of an OSCE-type assessment prepared in multiple, committed and competent hands may serve as a unique assessment and, in fact, assess the development of clinical skills for future doctors.

Keywords: *educational measurements, higher education, medical education, objective structured clinical examination*

Introduction

The teaching and learning process in higher education has progressed remarkably in recent years worldwide, including in medical schools (Irby et al., 2010; Mann, 2011). In direct alignment with this new pedagogical reality, Brazilian medical schools have been re-organizing their curricula by competencies and adapting to the implementation of new teaching strategies with different learning scenarios, incentives for research training and reassessment of evaluation methods (Santos, 2011).

In order to achieve all the points described in the National Guidelines (Conselho Nacional de Educação, 2014), the curriculum of medical courses should focus on making students able to develop more than theoretical skills during their training. Thus, professors must instruct future professionals committed to ethics, able to understand the role of social factors in maintaining health, the insertion of the individual in the family, their role in promoting health and adequate communication with the patients and their families. Therefore, adjustments to the knowledge pyramid suggested by Miller (1990) have been proposed, where, at the top of the pyramid, the student "consistently demonstrates the attitudes, values and behaviors expected of those who came to think, act and feel like a physician" (Cruess et al., 2016; Miller, 1990). Pedagogically, for this new professional model, changes in teaching strategies have become imperative. With more innovative curriculum design based in competencies, the adoption of active learning strategies have been recommended (Frohna et al., 2006; Prado et al., 2011; Schmidt et al., 2009).

Research Focus

Another crucial point is: how to evaluate students' performance in this new scenario? It is still widely accepted that assessment of academic performance is necessary to monitor student's development. Such assessment serves as a guide for assessing individual progress in incorporating the knowledge and skills necessary for medical profession. However, choosing an appropriate and efficient assessment method is not an easy task and remains debated (Lurie et al., 2009). Traditional approaches using cognitive assessment provide an inaccurate, incomplete picture of the holistic training, not being in alignment with the guidelines of the new methodologies proposed in competency-based curricula (Troncon, 2004). A more complete assessment should do more than measure the volume of correct information, it should be able to provide the desired evidence that students are achieving curricular objectives and competencies.

Of note, there are different types of assessment approaches and professors should always take into consideration which one is the most appropriate to assess whether the objectives of a given course and/or discipline have been achieved by students. The Structured Objective Clinical Examination (OSCE) is one of these assessment tools already successfully tested in different courses, including medical schools (Carraccio & Englander, 2000; Fischer et al., 2017; Kolivand et al., 2020). It is designed to test, in a practical way and in a controlled environment, the performance of clinical and main medical skills. Although it has already been tested and applied at major universities in the world such as in Canada, Germany, United Kingdom and the United States (Chisnall et al., 2015; Furmedge et al., 2016; Goldhamer et al., 2018; Lafleur et al., 2015; Müller et al., 2019; Stansfield et al., 2016), it is still not widely applied in Brazilian Medical schools in detriment of traditional assessments (Troncon, 2004).

In this context, it is interesting to note that recently in Brazil, some government programs in the areas of health and education have provided an explosive growth in the number of medical

courses. In these new medical schools, in the face of a new society driven by technology, teaching strategies have been rethought and adequate (Farias et al., 2015). Although, these schools proposed mostly innovative curricula based on interdisciplinarity and multi-professionalism with the use of problem-based learning (PBL) and Student-centered learning (Meireles et al., 2019; Takenami et al., 2018), the OSCE-type evaluation methods are still not widely used.

Literature Review

The Objective Structured Clinical Exam was initially proposed in the 1970s to measure the achievement of clinical skills (Harden et al., 1975). Since then, it has been applied (with variations) in an increasing number of medical schools worldwide (Carraccio & Englander, 2000; Patrício et al., 2013). In this scenario, this type of assessment and its variations have already been tested in different health courses, from basic sciences such as physiology and anatomy to more advanced units in the clinical axis such as urology, gynecology and even radiology (Dumont et al., 2016; Noureldin et al., 2015; Van Den Berk et al., 2011; Wani, 2015; Yaqinuddin et al., 2013).

The organization of the OSCE may vary between the institutions that apply it as an evaluation method, however the principles that underlie it are common among schools. In summary, students perform sequentially a series of cases and clinical stations structured to measure the achievement of skills such as (i) communication and professionalism, (ii) obtaining a history, (iii) physical examination, (iv) clinical reasoning and (v) practical / technical skills. Different strategies can be used in the stations, for example, the interaction with a simulated patient, a computer simulation, a series of clinical information etc. Each of these stations has a professor to assess student performance based on a pre-defined checklist. Studies show that if there is a compromise in the elaboration of the OSCE, it becomes a workable, valid, reliable, cheap and efficient method (Gormley, 2011; Zayyan, 2011).

It is important to carry out a critical analysis of this evaluation strategy. There is still a great deal of heterogeneity between models from different medical schools and even among the evaluators that participate of the evaluation process (Guraya et al., 2010; Schleicher et al., 2017). In addition, the experimental designs of the studies that evaluate the efficiency of the OSCE, in general, have important methodological limitations which make us take a cautious look at the subject which certainly needs a better exploration (Turner & Dankoski, 2008).

A point that must be highlighted, however, is the positive evaluation on the part of the students who consider the method efficient and perceive its advantages, despite finding it very stressful (Khorashad et al., 2014). Another important point is that the OSCE has already been tested in comparison with other assessment models demonstrating inconclusive levels of efficiency and perceptions (Chisnall et al., 2015; Lakshmiathy, 2015; Turner & Dankoski, 2008).

Research Aim

In this context, in order to contribute to the OSCE's validation as an effective evaluation model, it is essential to compare it with the other assessment models. In the present research, in students from a Brazilian private medical school, it was assessed the correlation of evaluation scores originated from different assessment methods commonly used in medical schools with scores from OSCE.

Research Methodology

General Background

A Cross-sectional, descriptive and analytical study was conducted with medical students in order to assess the correlation between their performances in OSCE and other types of evaluation tests. For this purpose, their scores in five different curricular areas of internship were analyzed.

Sample

One hundred ninety-eight students of the 5th year of Medicine course (Internship), in a private medical school from Salvador, Northeast Brazil. The analyzed curricular areas were “Pediatrics”, “Surgery”, “Internal Medicine”, “Family and Community Medicine” and “Obstetrics and Gynecology”. In this medical school, each student completes a two-month internship in each of the aforementioned areas developing activities at a training hospital, outpatient clinic, emergency unit, surgical and obstetric center, and Family Health Units. The variables used for analyses were epidemiologic data and grades of each student in distinct evaluations described below which were performed at the end of each rotation. The grades could range from 0 to 10 and have 1 decimal place.

Procedures and Data Collection

In order to test the cognitive domain, a theoretical and objective test was applied with 20 multiple-choice questions, dealing with the most common subjects in each area of medical knowledge. The main professor of each curricular unit prepared the test.

To test the ability skills, it was applied the Objective Structured Practical Exam (OSPE) model on low complexity simulators for the curricular units of Surgery and Pediatrics, preserving patient safety. Invasive, noninvasive procedures and clinical situations, analysis of laboratory and imaging tests were tested. The test was applied in the Simulation Center and graded each student according to the standardized form prepared by professors of the respective areas. For the curricular units of Internal Medicine, Family and Community Medicine, and Obstetrics and Gynecology, we used the mini-clinical evaluation exercise (Mini-Cex) (Megale et al., 2009).

With the purpose of test, the behavioral skills, the preceptor of each internship field performed a conceptual evaluation, following a standardized form, representing the student's performance during the 2-month rotation. The final grade to each student was the weighted average of 30% performance in cognitive assessment, 30% performance in abilities skills and 40% performance in the conceptual evaluation.

At the end of the internship, an OSCE was carried out among all the students. It was composed by stations of clinical attendance with actors, representing habitual clinical situations in Clinical Medicine, Family and Community Medicine and Surgical Clinic, a station of case-long Pediatrics and a video station on ethical conduct of diagnosis and information to the relatives of HIV patients. Professors specialized in each area of knowledge, graded each student, following a standardized single file.

All students, enrolled in the 5th year of the Medical course, performed all the evaluations and signed the Free and Informed Consent Form.

Statistical Analysis

The normality of the sample was evaluated by the Kolmogorov-Smirnov test (“Kolmogorov–Smirnov Test,” 2008). The mean values with standard deviation or median with interquartile range were used as measures of central tendency and dispersion, depending if a given variable presented or not Gaussian distribution in the study population. The comparison between the average grades of the evaluations was done through the Kruskal Wallis test with multiple Dunn comparisons or nonparametric linear trend post-test (Kruskal & Wallis, 1952).

To test the correlation between the assessments, it was used the Spearman correlation coefficient (Spearman, 2010). The p -values were considered statistically significant when $p < .05$. The OSCE assessment grade was considered the gold standard. This metric was used to study its correlation with the grades of the theoretical test (assessing cognitive competence), the marks of the practical test (assessing the abilities skills) and the grades of the conceptual evaluation (assessing the behavioral competence).

Ethical Considerations

The investigations were conducted according to the principles expressed in the Declaration of Helsinki (World Medical Association, 2013). This work was executed in accordance with Resolution no. 466/12 of the Brazilian National Health Council and was approved by the Institutional Review Board of Salvador University (UNIFACS, Laureate International Universities). The students who accepted to participate in the research signed Terms of Free and Clarified Compromise.

Research Results

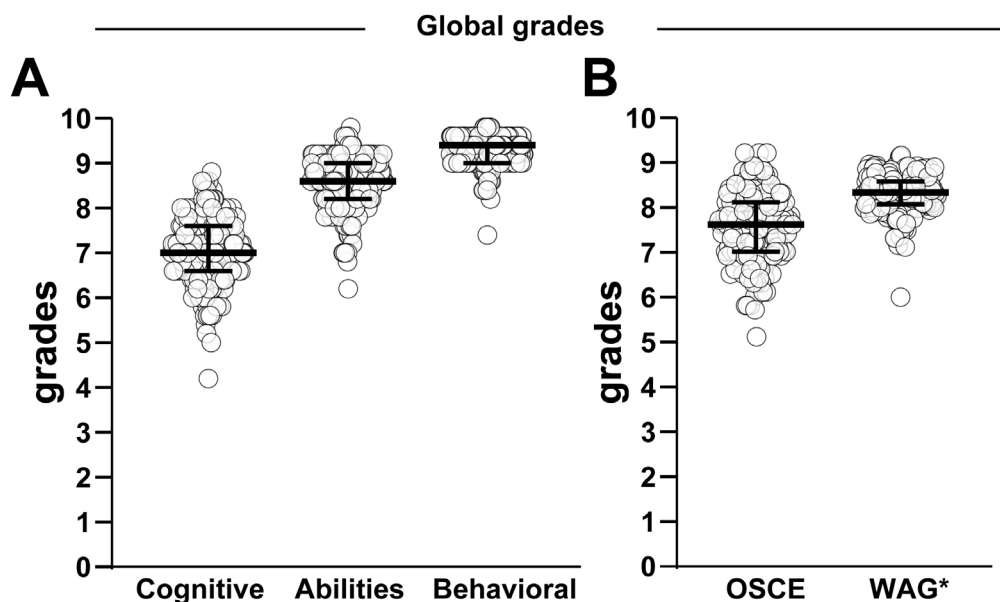
Descriptive Results

Results of 198 students were analyzed. The group comprised 144 female students (73%), with a mean age of 25.8 ± 3 years old, of high economic standard. Eighty percent (80%) of the students attended their first higher education course, but 20% had started or completed some other course previously, mostly in health.

The average final grade of the students in each of the 5 curricular units (global cognitive, global skills and overall behavioral) and OSCE were obtained: the Global cognitive mean of 7.0 ± 0.7 and the median 7.0; the overall global skills were 8.6 ± 0.6 and the median 8.6; the overall behavioral mean was 9.3 ± 0.3 and the median 9.4; the mean of the weighted average grade (WAG) was 8.3 ± 0.4 and the median 8.4; the mean of OSCE was 7.6 ± 0.8 and the median 7.7. The scatter plots of the grades are shown in Figure 1.

Figure 1

Global grades of students. In (A) the mean of cognitive, abilities and behavioral grades are presented. In (B) the OSCE and WAG grades are shown

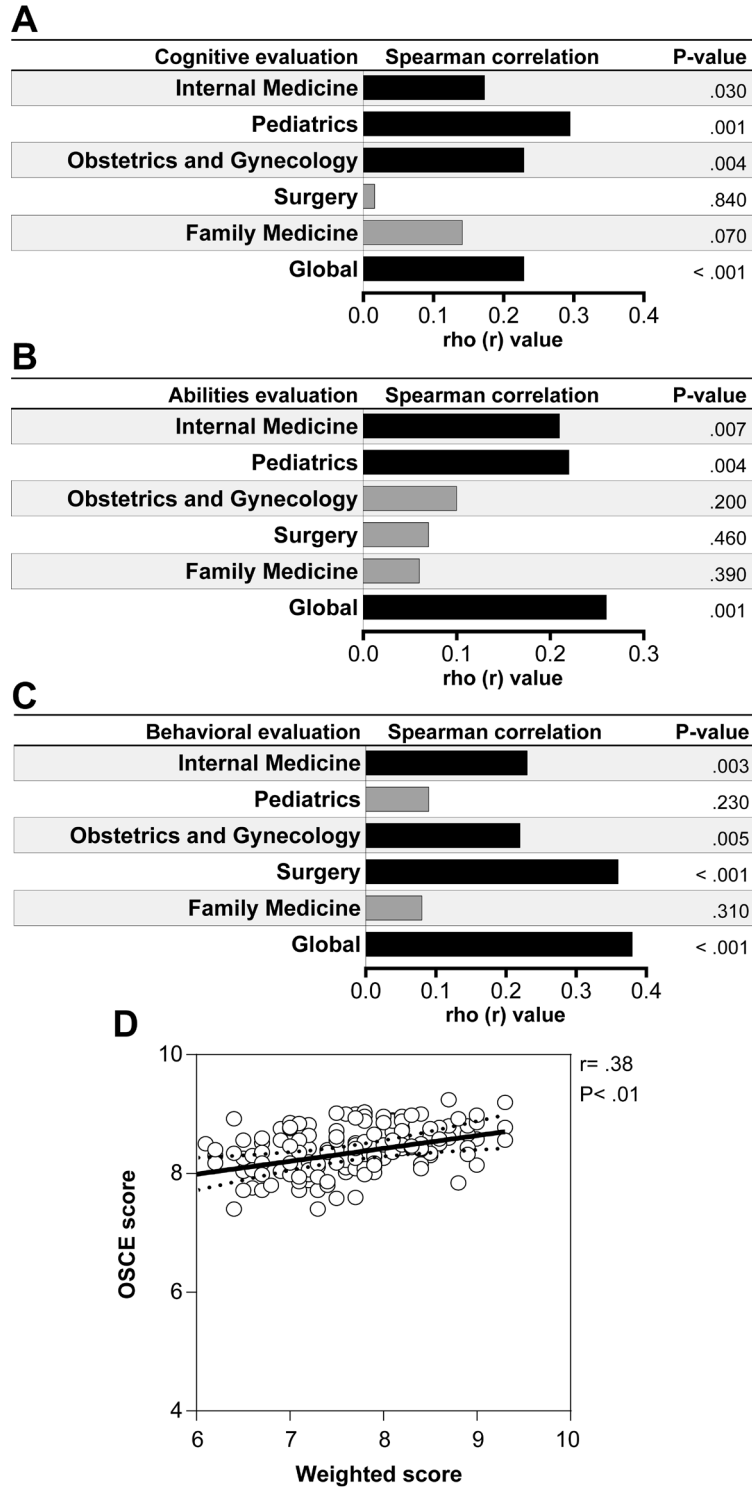


*WAG (weighted average grade) = (cognitive 0,3) + (ability x 0,3) + (behavioral x 0,4)

Analytical Results

To attest to the student's cognitive domain, it was used the traditional multiple-choice test. A weak positive but statistically significant correlation with the OSCE results in the overall score was detected ($r = .22$ $p < .001$). When analyzing the evaluations performed in each rotation, it was observed a correlation in Internal Medicine, Obstetrics and Gynecology and Pediatrics (Figure 2A). To test the skill domain, a practical test was applied. It was found a weak and statistically significant correlation with the OSCE results in the overall score with $r = .26$ and $p < .001$. When analyzing the evaluations performed in each rotation, a positive correlation in Internal Medicine and Pediatrics was found (Figure 2B). To attest the behavioral domain of the student, a conceptual evaluation was performed. It was found a weak and statistically significant positive correlation with the OSCE results in the overall score with $r = .38$ and $p < .001$. When analyzing the evaluations performed at each rotation, correlation in Internal Medicine, Obstetrics and Gynecology and Surgery were found (Figure 2C). Considering the final weighted score, we found a strong positive correlation with the OSCE $r = .38$ and $p < .001$ (Figure 2D).

Figure 2
Distribution per curricular unit of Spearman correlation vs OSCE. The Spearman correlation of (A) Cognitive, (B) Abilities, (C) Behavioral and (D) Weighted final grade evaluations with OSCE are presented



Discussion

In this research, the average age of the students was 25.8 ± 3 years old, which puts them in the sociological category of Generation Y. In addition, students mostly belong to a higher social class. Such characteristics make them widely familiar with technological resources and Internet, however, this generation of students is known for their impatience, immediacy and difficulty in following conventional models of work and hierarchy (Desy et al., 2017; Nalliah & Allareddy, 2014). Pedagogically, these characteristics must be considered in the teaching-learning and evaluation process.

It has been consistently demonstrated that this generation of students adapts well to the use of active learning methodologies (Quintanilha et al., 2018). In general, the use of more dynamic assessment tools results in better performance compared to more traditional tools. In the present study, it was found a lower performance in traditional cognitive assessments than in skill (practical or Mini-Cex) and conceptual assessments (Figure 1). This corroborates the findings of Sandoval, who evaluated the students' performance in the evaluations of the Pediatric Internship. The researcher detected approximately 20% lower scores on written exams than the result of observing daily clinical practice (Sandoval et al., 2010). A curriculum based on active methodologies may play an important role in these findings, as it stimulates students' practical and attitudinal character.

Here, we can clearly see a narrowing in the distribution of grades for conceptual assessment. Despite the use of a standardized instrument and the incentive to individualize the concept of each student, the values show low dispersion, with low sensitivity and the ability to detect low student performance. It demonstrates that these kinds of evaluations sometimes do not correspond to the students' real performance and it is important to think deeply about how to assess these cases better.

An important observation to be highlighted is that most health science professors working in the medical courses were not prepared for the teaching activities, reproducing, in general, the traditional model in which they were trained. Although institutions may offer update courses and pedagogical forums, most medical professors are unable to attend them (Quintanilha et al., 2020). This reality is observed in other studies (Costa, 2010) contributing to the maintenance of the wrong perception that the evaluation is only for approval (Diedrichs et al., 2012). They do not consider the main aim to correct the course according to the educational process and to act as an important pillar of the teaching-learning process (Furmedge et al., 2016). In this context, the continuous development of the professor can promote a paradigm shift, and this is extremely important for the construction of a reflective, positive and transformative evaluation philosophy.

Moreover, there is a gap between the recommendation of conducting classes using active learning methodologies and the common application of exams using only cognitive tests. The implementation of new assessment tools finds it difficult to convince professors about the validity of their use since, as previously discussed, they were trained and evaluated in traditional ways during their academic life. They complain about the number of checklists and bureaucratic formats that competency-based assessment models establish (Lombarts, 2015). However, based on continuous practice and the consequent perception of the extension of this type of evaluation, there is a decrease in resistance and such changes in evaluation methods can prove to be of great value in the process of medical training based on competences and in the training of new future physicians.

In this sense, studies show that the OSCE presents itself as a complete and reliable form of assessment in medical schools (Matsuo et al., 2013; Townsend et al., 2001; Walters et al., 2005; Zayyan, 2011). In the present research, it was demonstrated that the student's performance in the various assessments can be proportional to their performances in the OSCE, with a statistically

significant correlation. Our results corroborate the data from the study of Couto (2019) who demonstrated that assessment strategies in problem-based learning (PBL) approach, including Formative Assessment, Progress Testing and OSCE have positive correlations, which increases as the medical course becomes more complex (Couto et al., 2019).

In this regard, the OSCE, despite the greater complexity of its execution, especially in large groups, must be evaluated by medical education institutions as it is a comprehensive assessment that effectively measures various degrees of medical skills. This assessment modality can even replace a huge range of assessments from different disciplines which can have their main contents and competencies covered at once in a single exam.

Conclusions and Implications

The definition of efficient forms of assessment is an issue widely discussed in the field of education. In Medical education, with the curricular changes that have occurred in the last few years, such as the introduction of models such as Problem Based Learning and Active Learning strategies, it is incoherent to maintain the same traditional forms of assessment. The Objective Structured Clinical Exam (OSCE) has been presented as a useful and effective alternative in several Medical schools. In the present study, we compared the performance of students who undertook different assessment strategies with their OSCE performance. It was observed that there is a weak or moderately positive relationship with statistical relevance between them.

In practical terms, considering the OSCE's correlation with other assessment models and especially the large number of tests that students usually take and the great variability among them, the proposal for the organization of an OSCE-type assessment prepared in multiple, committed and competent hands, can serve as a unique assessment and, in fact, assess the development - or not - of clinical skills for future physicians.

Since we are evaluating a single group of students from a specific university, the results presented may not be extrapolated to other educational contexts. However, there are many studies that corroborate with our data making us believe that, in fact, our results are representative and suggest a reality in medical courses in general.

So, there is an urgent need for new, robust and analytical studies that assess the effectiveness and perception of learning, as well as application models that increase quality and decrease heterogeneity and bias. In this way, medical education educators and managers will be able to base their practices on what we call evidence-based medical education.

Limitations and Suggestions

As this is an observational study, the results obtained are influenced by the subjectivity of the evaluation process, from the perspective of each professor and do not allow us to conclude casual relationships. An interesting suggestion is training professors in the standardization of the checklists and cognitive tests, in addition to training of each professor that teaches in practical setting the importance of the objectivity on behavioral assessments, to analyze whether the evaluations make it even closer to the “gold-standard” assessment.

Acknowledgements

The authors would like to thank Ms. Galyna Blyznyuk for her assistance in revising the English text.

References

- Carraccio, C., & Englander, R. (2000). The objective structured clinical examination. *Archives of Pediatrics & Adolescent Medicine*, 154(7), 736. <https://doi.org/10.1001/archpedi.154.7.736>
- Chisnall, B., Vince, T., Hall, S., & Tribe, R. (2015). Evaluation of outcomes of a formative objective structured clinical examination for second-year UK medical students. *International Journal of Medical Education*, 6, 76–83. <https://doi.org/10.5116/ijme.5572.a534>
- Conselho Nacional de Educação. (2014, June 3). *Institui diretrizes curriculares nacionais do curso de graduação em medicina e dá outras providências* [Institutes national curricular guidelines for the undergraduate course in medicine and provides other measures]. Ministério Da Educação. http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=15874-rces003-14&category_slug=junho-2014-pdf&Itemid=30192
- Costa, N. M. S. C. (2010). Pedagogical training of medicine professors. *Revista Latino-Americana de Enfermagem*, 18(1), 102–108. <https://doi.org/10.1590/s0104-11692010000100016>
- Couto, L. B., Durand, M. T., Wolff, A. C. D., Restini, C. B. A., Faria, M., Romão, G. S., & Bestetti, R. B. (2019). Formative assessment scores in tutorial sessions correlates with OSCE and progress testing scores in a PBL medical curriculum. *Medical Education Online*, 24(1), Article 1560862. <https://doi.org/10.1080/10872981.2018.1560862>
- Cruess, R. L., Cruess, S. R., & Steinert, Y. (2016). Amending Miller's pyramid to include professional identity formation. *Academic Medicine*, 91(2), 180–185. <https://doi.org/10.1097/acm.0000000000000913>
- Desy, J. R., Reed, D. A., & Wolanskyj, A. P. (2017). Milestones and millennials: A perfect pairing—competency-based medical education and the learning preferences of generation Y. *Mayo Clinic Proceedings*, 92(2), 243–250. <https://doi.org/10.1016/j.mayocp.2016.10.026>
- Diedrichs, P., Willen Lundgren, B., & Karlsudd, P. (2012). Flexible examination as a pathway to learning. *Problems of Education in the 21st Century*, 40, 26–34. http://www.scientiasocialis.lt/pec/node/files/pdf/vol40/26-34.Diedrichs_Vol.40.pdf
- Dumont, T., Hakim, J., Black, A., & Fleming, N. (2016). Does an advanced pelvic simulation curriculum improve resident performance on a pediatric and adolescent gynecology focused objective structured clinical examination? A cohort study *Journal of Pediatric and Adolescent Gynecology*, 29(3), 276–279. <https://doi.org/10.1016/j.jpjag.2015.10.015>
- Farias, P. A. M., Martin, A. L. A. R., & Cristo, C. S. (2015). Aprendizagem ativa na educação em saúde: Percurso histórico e aplicações [Active learning in health education: Historical background and applications]. *Revista Brasileira de Educação Médica*, 39(1), 143–150. <https://doi.org/10.1590/1981-52712015v39n1e00602014>
- Fischer, M. A., Kennedy, K. M., Durning, S., Schijven, M. P., Ker, J., O'Connor, P., Doherty, E., & Kropmans, T. J. B. (2017). Situational awareness within objective structured clinical examination stations in undergraduate medical training - a literature search. *BMC Medical Education*, 17(1), 1–10. <https://doi.org/10.1186/s12909-017-1105-y>
- Frohna, A. Z., Hamstra, S. J., Mullan, P. B., & Gruppen, L. D. (2006). Teaching medical education principles and methods to faculty using an active learning approach: The university of Michigan medical education scholars' program. *Academic Medicine*, 81(11), 975–978. <https://doi.org/10.1097/01.acm.0000242573.71314.74>
- Furmedge, D. S., Smith, L.-J., & Sturrock, A. (2016). Developing doctors: What are the attitudes and perceptions of year 1 and 2 medical students towards a new integrated formative objective structured clinical examination? *BMC Medical Education*, 16(1), 1–9. <https://doi.org/10.1186/s12909-016-0542-3>
- Goldhamer, M. E. J., Cohen, A., Brooks, M., Macklin, E. A., Co, J. P. T., & Weinstein, D. (2017). Use of an objective structured clinical exam (OSCE) for early identification of communication skills deficits in interns. *Medical Teacher*, 40(1), 40–44. <https://doi.org/10.1080/0142159x.2017.1387646>
- Gormley, G. (2011). Summative OSCEs in undergraduate medical education. *Ulster Medical Journal*, 80(3), 127–132. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3605523/pdf/umj0080-0127.pdf>

- Guraya, S. Y., Alzobydi, A. H., & Salman, S. (2010). Objective structured clinical examination: Examiners' bias and recommendations to improve its reliability. *Journal of Medicine and Medical Science, 1*(7), 269–272. <https://www.interestjournals.org/medicine-medical-sciences.html>
- Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective structured examination. *BMJ, 1*(5955), 447–451. <https://doi.org/10.1136/bmj.1.5955.447>
- Irby, D. M., Cooke, M., & O'Brien, B. C. (2010). Calls for reform of medical education by the Carnegie Foundation for the advancement of teaching: 1910 and 2010. *Academic Medicine, 85*(2), 220–227. <https://doi.org/10.1097/acm.0b013e3181c88449>
- Kolivand, M., Esfandyari, M., & Heydarpour, S. (2020). Examining validity and reliability of objective structured clinical examination for evaluation of clinical skills of midwifery undergraduate students: A descriptive study. *BMC Medical Education, 20*(1), 1–7. <https://doi.org/10.1186/s12909-020-02017-4>
- Kolmogorov–Smirnov Test. (2008). *The concise encyclopedia of statistics. Springer*, 283–287. https://doi.org/10.1007/978-0-387-32833-1_214
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621. <https://www.jstor.org/stable/2280779?origin=JSTOR-pdf&seq=1>
- Lafleur, A., Côté, L., & Leppink, J. (2015). Influences of OSCE design on students' diagnostic reasoning. *Medical Education, 49*(2), 203–214. <https://doi.org/10.1111/medu.12635>
- Lakshmiopathy, K. (2015). MBBS student perceptions about physiology subject teaching and objective structured practical examination based formative assessment for improving competencies. *Advances in Physiology Education, 39*(3), 198–204. <https://doi.org/10.1152/advan.00073.2014>
- Lombarts, M. J. M. H. (2015). Competence-based education misses the essence of the medical profession. *Perspectives on Medical Education, 4*(6), 326–328. <https://doi.org/10.1007/s40037-015-0233-5>
- Lurie, S. J., Mooney, C. J., & Lyness, J. M. (2009). Measurement of the general competencies of the accreditation council for graduate medical education: A systematic review. *Academic Medicine, 84*(3), 301–309. <https://doi.org/10.1097/acm.0b013e3181971f08>
- Mann, K. V. (2010). Theoretical perspectives in medical education: Past experience and future possibilities. *Medical Education, 45*(1), 60–68. <https://doi.org/10.1111/j.1365-2923.2010.03757.x>
- Matsuo, O., Matsuo, K., Kawao, N., Tamura, Y., Yano, M., Okada, K., Nagai, N., & Ueshima, S. (2013). Development of an objective structured skill examination in physiology education. *Problems of Education in the 21st Century, 52*, 1–8. http://www.scientiasocialis.lt/pec/node/files/pdf/vol52/84-91.Matsuo_Vol.52.pdf
- Megale, L., Gontijo, E. D., & Motta, J. A. C. (2009). Avaliação de competência clínica em estudantes de medicina pelo Miniexercício Clínico Avaliativo (Miniex) [Evaluation of medical students' clinical skills using the Mini-Clinical Evaluation Exercise (mini-CEX)]. *Revista Brasileira de Educação Médica, 33*(2), 166–175. <https://doi.org/10.1590/s0100-55022009000200002>
- Meireles, M. A. C., Fernandes, C. C. P., & Silva, L. S. (2019). Novas diretrizes curriculares nacionais e a formação médica: Expectativas dos discentes do primeiro ano do curso de medicina de uma instituição de ensino superior [New national curricular guidelines and medical education: Expectations of first year students of the medical course of a higher education institution]. *Revista Brasileira de Educação Médica, 43*(2), 67–78. <https://doi.org/10.1590/1981-52712015v43n2rb20180178>
- Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63-7. <https://doi.org/10.1097/00001888-199009000-00045>
- Müller, S., Koch, I., Settmacher, U., & Dahmen, U. (2019). How the introduction of OSCEs has affected the time students spend studying: Results of a nationwide study. *BMC Medical Education, 19*(1), 1–7. <https://doi.org/10.1186/s12909-019-1570-6>
- Nalliah, R. P., & Allareddy, V. (2014). Weakest students benefit most from a customized educational experience for Generation Y students. *PeerJ, 2*, 1–11. <https://doi.org/10.7717/peerj.682>
- Noureddin, Y. A., Elkoushy, M. A., Fahmy, N., Carrier, S., Elhilali, M. M., & Andonian, S. (2015). Assessment of photoselective vaporization of prostate skills during Urology Objective Structured Clinical Examinations (OSCE). *Canadian Urological Association Journal, 9*(1–2), 61. <https://doi.org/10.5489/cuaj.2273>

- Patrício, M. F., Julião, M., Fareleira, F., & Carneiro, A. V. (2013). Is the OSCE a feasible tool to assess competencies in undergraduate medical education? *Medical Teacher*, 35(6), 503–514. <https://doi.org/10.3109/0142159x.2013.774330>
- Prado, H. M., Hannois Falbo, G., Rodrigues Falbo, A., & Natal Figueirôa, J. (2011). Active learning on the ward: Outcomes from a comparative trial with traditional methods. *Medical Education*, 45(3), 273–279. <https://doi.org/10.1111/j.1365-2923.2010.03846.x>
- Quintanilha, L. F., Costa, G. N., & Coutinho, M. R. (2018). Medical student perceptions about active methodologies in the study of physiology in medical schools in Salvador, Brazil. *Advances in Physiology Education*, 42(4), 693–696. <https://doi.org/10.1152/advan.00105.2018>
- Quintanilha, L. F., Santos Silva Farias, C., & Andrade, B. B. (2020). Training and pedagogical involvement among higher education professors in health: An analysis of medical courses. *Revista Internacional de Educação Superior*, 7, 1–17. <https://periodicos.sbu.unicamp.br/ojs/index.php/riesup/article/view/8659850/22697>
- Sandoval, G. E., Valenzuela, P. M., Monge, M. M., Toso, P. A., Trivino, X. C., Wright, A. C., Paris, E., Sanchez, I., & Valdivia, G. S. (2010). Analysis of a learning assessment system for pediatric internship based upon objective structured clinical examination, clinical practice observation and written examination. *Jornal de Pediatria*, 86(2), 131–136. <https://doi.org/10.2223/jped.1986>
- Santos, W. S. (2011). Organização curricular baseada em competência na educação médica [Competency-based curriculum organization in medical education]. *Revista Brasileira de Educação Médica*, 35(1), 86–92. <https://doi.org/10.1590/s0100-55022011000100012>
- Schleicher, I., Leitner, K., Juenger, J., Moeltner, A., Ruesseler, M., Bender, B., Sterz, J., Schuettler, K.-F., Koenig, S., & Kreuder, J. G. (2017). Examiner effect on the objective structured clinical exam – a study at five medical schools. *BMC Medical Education*, 17(1), 1–7. <https://doi.org/10.1186/s12909-017-0908-1>
- Schmidt, H. G., Cohen-Schotanus, J., & Arends, L. R. (2009). Impact of problem-based, active learning on graduation rates for 10 generations of Dutch medical students. *Medical Education*, 43(3), 211–218. <https://doi.org/10.1111/j.1365-2923.2008.03287.x>
- Spearman, C. (2010). The proof and measurement of association between two things. *International Journal of Epidemiology*, 39(5), 1137–1150. <https://doi.org/10.1093/ije/dyq191>
- Stansfield, R. B., Diponio, L., Craig, C., Zeller, J., Chadd, E., Miller, J., & Monrad, S. (2016). Assessing musculoskeletal examination skills and diagnostic reasoning of 4th year medical students using a novel objective structured clinical exam. *BMC Medical Education*, 16(1), 1–7. <https://doi.org/10.1186/s12909-016-0780-4>
- Takenami, I. O., Palácio, M. A. V., Andrade, W., & Cansanção, I. F. (2018). Use of active learning methodologies in medical teaching institutions in northeastern Brazil. *Revista de Educação Da Universidade Federal Do Vale Do São Francisco*, 8(17), 1–17. <https://www.periodicos.univasf.edu.br/index.php/revasf/article/view/239>
- Townsend, A. H., McIlvenny, S., Miller, C. J., & Dunn, E. V. (2001). The use of an objective structured clinical examination (OSCE) for formative and summative assessment in a general practice clinical attachment and its relationship to final medical school examination performance. *Medical Education*, 35(9), 841–846. <https://doi.org/10.1046/j.1365-2923.2001.00957.x>
- Troncon, L. E. A. (2004). Clinical skills assessment: Limitations to the introduction of an “OSCE” (Objective Structured Clinical Examination) in a traditional Brazilian medical school. *Sao Paulo Medical Journal*, 122(1), 12–17. <https://doi.org/10.1590/s1516-31802004000100004>
- Turner, J. L., & Dankoski, M. E. (2008). Objective structured clinical exams: A critical review. *Family Medicine*, 40(8), 574–578. <https://pubmed.ncbi.nlm.nih.gov/18988044/>
- Van den Berk, I. A. H., van de Ridder, J. M. M., & van Schaik, J. P. J. (2011). Radiology as part of an objective structured clinical examination on clinical skills. *European Journal of Radiology*, 78(3), 363–367. <https://doi.org/10.1016/j.ejrad.2008.10.042>
- Walters, K., Osborn, D., & Raven, P. (2005). The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. *Medical Education*, 39(3), 292–298. <https://doi.org/10.1111/j.1365-2929.2005.02091.x>
- Wani, P. (2015). Traditional clinical examination vs objective structured practical examination in human physiology: Examiners bias. *International Journal of Medical Science and Public Health*, 4(5), 607. <https://doi.org/10.5455/ijmsph.2015.17042014122>

- World Medical Association Declaration of Helsinki. (2013). *JAMA*, 310(20), 2191. <https://doi.org/10.1001/jama.2013.281053>
- Yaqinuddin, A., Zafar, M., Ikram, M. F., & Ganguly, P. (2012). What is an objective structured practical examination in anatomy? *Anatomical Sciences Education*, 6(2), 125–133. <https://doi.org/10.1002/ase.1305>
- Zayyan, M. (2011). Objective structured clinical examination: The assessment of choice. *Oman Medical Journal*, 26(4), 219–222. <https://doi.org/10.5001/omj.2011.55>

Received: *June 03, 2020*

Accepted: *September 12, 2020*

Cite as: Castellani, L., Quintanilha, L. F., Arriaga, M. B., de Lourdes Lima, M., & Andrade, B. B. (2020). Objective structured clinical examination (OSCE) as a reliable evaluation strategy: Evidence from a Brazilian medical school. *Problems of Education in the 21st Century*, 78(5), 674-687. <https://doi.org/10.33225/pec/20.78.674>

Luciana Castellani	PhD, Bahiana School of Medicine and Public Health (Escola Bahiana de Medicina e Saúde Pública - EBMSp), R. Silveira Martins, 100 - Cabula, Salvador, Bahia, Brazil. Professor, Salvador University (Universidade Salvador – UNIFACS), R. Dr. José Peroba, 251 - Stiep, Salvador, Bahia, Brazil. E-mail: lucianacastellani@gmail.com ORCID: https://orcid.org/0000-0002-5000-9244
Luiz Fernando Quintanilha (Corresponding author)	Professor, Salvador University (Universidade Salvador - UNIFACS), R. Dr. José Peroba, 251 - Stiep, Salvador, Bahia, Brazil. Professor, FTC University Center (UniFTC), Av. Luís Viana, 8812 - Paralela, Salvador, Bahia, Brazil. E-mail: quintanilha.lf@gmail.com ORCID: https://orcid.org/0000-0001-8911-9806
Maria Belen Arriaga	Researcher, Oswaldo Cruz Foundation (Fundação Oswaldo Cruz - FIOCRUZ), R. Waldemar Falcão, 121 - Candeal, Salvador, Bahia, Brazil. E-mail: tmbag711@gmail.com ORCID: https://orcid.org/0000-0001-6883-8422
Maria de Lourdes Lima	Professor, Bahiana School of Medicine and Public Health (Escola Bahiana de Medicina e Saúde Pública - EBMSp), R. Silveira Martins, 100 - Cabula, Salvador, Bahia, Brazil. Professor, Salvador University (Universidade Salvador – UNIFACS), R. Dr. José Peroba, 251 - Stiep, Salvador, Bahia, Brazil. E-mail: mlourdeslima@bahiana.edu.br ORCID: http://orcid.org/0000-0002-2081-4162
Bruno Bezerril Andrade	Researcher, Oswaldo Cruz Foundation (Fundação Oswaldo Cruz - FIOCRUZ), R. Waldemar Falcão, 121 - Candeal, Salvador, Bahia, Brazil. Professor, Bahiana School of Medicine and Public Health (Escola Bahiana de Medicina e Saúde Pública - EBMSp), R. Silveira Martins, 100 - Cabula, Salvador, Bahia, Brazil. E-mail: bruno.andrade@fiocruz.br ORCID: https://orcid.org/0000-0001-6833-3811