

Adaptation, psychometric properties and implementation of the Mini-CEX in dental clerkship.

Claudia P. Véliz. DDS, MPH.

Assistant Professor. School of Dentistry, Faculty of Medicine, Pontificia Universidad Católica de Chile.

Javiera Fuentes-Cimma. MMedEd.

Assistant Professor. Physiotherapy program, Department of Health Sciences. Faculty of Medicine, Pontificia Universidad Católica de Chile.

Eduardo Fuentes-López. PhD

Assistant Professor. Speech-language pathology program, Department of Health Sciences. Faculty of Medicine, Pontificia Universidad Católica de Chile.

Arnoldo Riquelme. MD, MMedEd

Full Professor, Department of Gastroenterology, Centre for Medical Education and Health Sciences, Department of Health Sciences. Faculty of Medicine, Pontificia Universidad Católica de Chile.

ABSTRACT

Background: Workplace-based assessment is a key component of dental students during their clerkship in order to demonstrate clinical proficiency. **Purpose:** This study aims to analyze the results obtained from adapting the Mini-Clinical Evaluation Exercise (Mini-CEX) to the clerkship of a dentistry program, and to examine its psychometric properties. **Methods:** First, a Delphi panel method was used to ensure content validity. Then, the Mini-CEX was piloted in the dental clerkship where each student was assessed by at least two supervisors and a peer student; the psychometric properties, acceptability, and observation-time were analyzed afterward. **Results:** The pilot study was conducted between July and November of 2019. One hundred forty Mini-CEX were carried out on 30 students, eighty-four by supervisors and fifty-six by peers. The adapted instrument was proved to be unidimensional, obtaining an acceptable internal consistency ($\alpha=0.74$). There was a difference in observation-time as the type of assessor changed; the medians (Q1-Q3) were 10 minutes (5-15) for supervisors and 30 minutes (20-45) for peer students ($p<0.001$). This difference was also observed when analyzing the assessor's satisfaction ($p<0.001$) since the supervisor scored a median of 6 (6-6.75), and peer students scored a median of 7 (6-7). No differences were found between the scores given by supervisors and peers. **Conclusion:** The adapted version of the Mini-CEX can objectively assess students' clinical performance of dental student based on values of validity and reliability, which are similar to those obtained in the original instrument.

Keywords: Medical Education, Workplace-based assessment, Educational assessment, Clinical Clerkship, Dental Education.

INTRODUCTION

Assessment of learning is a fundamental part of the educational process since it provides both teachers and students with relevant information on the achieved learning outcomes.¹ The paradigm shift from the assessment *of* learning to assessment *for* learning involves that evaluation opportunities become a source of learning since they are part of the educational program as low-stakes assessments and not only at the end as high-stakes.^{2,3} Consequently, this approach has arisen as an important topic within the literature on health sciences education, since higher-educational institutions must certify that students have acquired the required competencies and meet the minimum standards for professional practice. Regarding the different levels proposed by Miller, ranging from “knows”, “knows how”, “shows how” and “does”,⁴ a variety of instruments designed to assess each one of them are described in the literature.¹ The clerkships' biggest challenge has been to have valid instruments to assess the “does” level.

Clerkships training is the final supervised clinical undergraduate practice of health professions, where students need to demonstrate competencies in real contexts. Evidence suggests that systematic assessments of clinical encounters with patients should be conducted to assess their clinical performance.⁵ A large number of instruments have been designed to assess the students' performance in the clinical workplace [workplace-based assessment (WBA)]. This type of assessment includes instruments that assess: the discussion of clinical cases; multisource assessment; portfolio, and direct observation tools.⁶

The American Board of Internal Medicine developed the mini-clinical evaluation exercise (Mini-CEX) in 1995. John Norcini carried out the first research on the assessment of clinical competencies in the workplace using the Mini-CEX.⁷ During the assessment process, a student is observed by a assessor while interacting with a patient in a real clinical context.^{8,9} Normally, the encounter has a brief duration, so the Mini-CEX can be applied systematically and conducted by

several assessors during a clinical rotation.¹⁰ Its psychometric properties have been reported in a systematic review,¹⁰ which reported acceptable reliability, which improves with proper previous training, increasing inter-rater reliability.^{10,11}

Although the Mini-CEX was designed to be applied in medical residents,⁵ it has been adapted to be used in other health profession programs, such as dentistry,^{6,12} physiotherapy,^{13,14} pharmacy,¹⁵ nursing^{16,17} and in medicine clerkships,^{18,19} due to its simple and brief implementation.

Some studies that report the use of this instrument in undergraduate^{12,20} and postgraduate^{6,21–24} dental students. Those studies described some positive aspects like the possibility of multiple observations by different supervisors in diverse clinical circumstances,²³ the possibility of effective, timely individualized feedback²⁵ and the possibility of assessing the capacity of reflection, self-perception,^{20,25} communication skills, and professionalism.¹² However, there are still no studies focusing on validation or psychometric analysis of the Mini-CEX implementation in dental contexts. In this way, this study will provide a valid and reliable instrument, and the process described here could be used as a reference in the construction of similar educational measurement instruments focused on WBA.

The aims of the present study were to analyze the results obtained from adapting the Mini-CEX to the clerkship of a dentistry program, and to examine its psychometric properties.

MATERIAL AND METHOD

The Ethics Committee of the Faculty of Medicine at Pontificia Universidad Católica de Chile approved this study, project number 190531001.

Adaptation of Mini-CEX for implementation in dentistry clerkship

Firstly, the dental clerkship's learning outcomes were thoroughly reviewed to decide which ones were going to be assessed using the Mini-CEX. Secondly, the selected learning outcomes were compared to those in the original instrument by using a Spanish-translated version,²⁶ and a prototype was formulated, consisting of 8 items: 1) Anamnesis and physical examination; 2) Clinical judgement; 3) Execution of the treatment; 4) Planning and organization of the session; 6) Efficient use of resources; 6) Professionalism; 7) Communication; & 8) Global competence. Each item was described as a series of behaviors that must be observed in students and a scale of 1 to 7 was preferred to be used since it is more culturally acceptable than the original scale of 1 to 9.

Participants and application

This study was conducted in de Dentistry School of the Pontificia Universidad Católica de Chile. The adapted Mini-CEX was applied to dental clerks between July and November of 2019, on a 16-week clerkship clinical rotation, each student was assessed by at least two supervisors and a peer. The assessors were trained to conduct the assessment during the first week of the semester, the first assessment of a supervisor and a peer was conducted between the second and the seventh week of the semester, whereas the second assessment (of a supervisor and a peer) was conducted between the eighth and fifteenth week of the semester.

Sample size

The sample size was calculated for the validity study of the Mini-CEX scale. The validity of the Mini-CEX was obtained using a confirmatory factor analysis. Thus, to calculate the sample size, the method based on the confirmatory factor analysis goodness-of-fit index was used.²⁷ A root mean square error of approximation (RMSEA) test was used to obtain the sample size needed to test the confirmatory factorial model's close-fit hypothesis. In this approach, the RMSEA test value should be set at a level that reflects an acceptable-fit model (i.e., RMSEA between 0.05 and 0.08); thus, considering an RMSEA=0.08 with 80% power, with eight items and one factor of the Mini-CEX scale, and $\alpha=0.05$, the sample size required was 150 assessments.

Psychometric properties

Content validity

Content validity was determined with a Delphi panel with national coverage, using a Google form to collect data and a five-point Likert scale. It was defined that domains with a 4.5 or higher score would be included in the adapted instrument. For each round of the Delphi panel, thirty-six people participated in consultations (supervisors and experts in education, among others).²⁸

Construct Validity: Classical Test Theory

An exploratory and confirmatory factor analysis was performed. These multivariate analysis techniques belong in the Classical Test Theory. The exploratory factor analysis aimed to detect latent variables that underlie the observed variables (instrument items). The confirmatory factor analysis made it possible to confirm the number of existing latent factors and their nature in the collected data.

Construct Validity: Item Response Theory (IRT)

IRT models allow latent traits to be related to the probability of response to an item, and they have become widely used in psychometry and recently in health sciences.²⁹ Based on this model, each question's query difficulty was estimated with respect to the assessed latent factor. Furthermore, it made it possible to evaluate whether a higher level of the construct assessed in the Mini-CEX entails an increase in the probability of achieving a better performance in each one of the Mini-CEX items.

Internal consistency

To determine internal consistency, the reliability coefficient (Cronbach's alpha) was estimated. This analysis was carried out once the number of items included in the Mini-CEX was confirmed.³⁰ A cut-off point of 0.70 was considered to establish an acceptable internal consistency.³¹

Acceptability

The satisfaction level of the assessor and the student was measured on a scale of 1 to 7. Differences in their satisfaction levels were sought when the evaluators were supervisors or peers.

Application time and feedback time

Both the time for observing the student and the time needed for feedback were recorded. Differences were sought when the evaluators were supervisors or peers.

Statistical Analysis

An exploratory analysis of data was carried so that atypical values could be found. By means of the Shapiro-Wilk test, the distribution of the continuous quantitative variables was determined. Descriptive statistics were estimated using the mean for continuous variables with normal distribution, whereas for those with biased distribution, the median (Q1-Q3) was used. In the case of categorical variables, a relative and absolute frequency was obtained. By a nonparametric equality-of-medians test, the medians of the assessments conducted by supervisors and peers for each item of the Mini-CEX scale were compared. This same test was used to compare the satisfaction level, application time, and feedback time.

The psychometric properties were obtained through Classical Test Theory (exploratory and confirmatory factor analyses) and IRT. In the case of the exploratory factor analysis, the principal component method was applied, having standardized factor loadings as a result. The factor loadings represent the relationship between the latent factor (measured construct) and the observed performances for each item of the Mini-CEX. As to the confirmatory analysis, a model in which all domains loaded on a single factor (unidimensional instrument) was specified. The coefficient of determination (R^2) was obtained by quantifying the percent variance of the instrument domains, which is explained by the identified factor.

Based on IRT, a one-parameter logistic model (1PL) was specified, estimating each item's difficulty parameter. Although the Mini-CEX form differentiated three possible performances (insufficient, sufficient and superior), there was variability in most items of the sufficient and superior performances. Given this circumstance, all responses were changed to make them dichotomous. Afterward, characteristic curves of the items were obtained, depicting a graph where the instrument items with higher difficulty have a greater slope. These curves allowed to evaluate whether a higher level of the construct assessed entailed an increase in the probability of achieving a superior performance in each of the Mini-CEX questions.

All analyses were performed using the STATA version 16 software and Mplus version 7. In both cases, it was taken into account the fact that repeated measures were used from each student.

RESULTS

Participants and implementation

As a result of the adaptation of the Mini-CEX and validation of experts, an adapted Mini-CEX with the following 8 items was designed: 1) Anamnesis and physical examination focused on the patient's condition; 2) Clinical judgement; 3) Autonomous execution of the treatment according to level; 4) Planning and organization of the session; 5) Efficient use of resources; 6) Professionalism; 7) Communication with the work team and patient; & 8) Global competence.

Boxes are included to record the observation-time, the time needed for feedback, and a box through which a domain can be pointed to as "not observed." It is requested that the assessors and student register their satisfaction level on a scale of 1 to 7. Lastly, a blank space was left to write down the feedback. On the reverse side of the adapted Mini-CEX, all domains and observable behaviors of students were described (Figure 1).

The adapted Mini-CEX was administered in thirty clerkship students of the dentistry program. By the end of the semester, 140 assessments had been conducted, of which supervisors performed eighty-four, and fifty-six were performed by peers, having a median of 5 per student (4-5). Not all assessors completed the additional information requested.

Item one did not present any variation in its scores as all students achieved the maximum score. In most of the forms (93.6%), all items were given scores belonging to the "superior" range (6 or 7). Only one student obtained a score within the range "unsatisfactory" (1, 2, or 3). No differences were observed between the ratings made by supervisors or peers.

Psychometric properties

Content validity

Were obtained twenty-seven responses (75%) in the first round and twenty-two responses (61%) at the second round of panel Delphi. After the second round, an agreement was reached on all domains. The first-round participants' suggestions were taken into account to improve the domain descriptions in the second round. Only items that were not agreed upon or modified needed to be submitted for discussion in the second round.

Finally, descriptions of each of the domains and behaviors that assessors must observe at the moment of assessment were elaborated in conjunction with the supervisors (Table 1). This activity is a fundamental part of the supervisor's and student's training.

Construct validity: Classical Test Theory

In the exploratory factor analysis, a factor with an Eigenvalue greater than one was identified. The scree plot showed an inflection point in the slope from the first factor (Figure 2). According to this, the adapted Mini-CEX is a unidimensional instrument. This factor explained 77.3% of the variability of responses.

All standardized factor loadings were high (>0.3), except for items five and seven (Table 3). The remaining loadings fluctuated between 0.34 and 0.93 for items six and two, respectively. Regarding the confirmatory factor analysis, all standardized factor loadings obtained values of 0.3 or higher (Table 3), except for items five and seven. Moreover, there are predictions with R^2 values of less than 0.1 (items five and seven), whereas others scored 0.87 (item two). It means that at least 87% of the variability of item two' scores was explained by the single factor identified.

Construct validity: Item Response Theory (IRT)

The assumption of unidimensionality (a single factor identified) of IRT models was met. As to the characteristic curves of the item, it was observed that the instrument items have a similar

difficulty as they showed a similar slope (Figure 3). Likewise, Figure 3 shows that the higher the level of the construct assessed by the Mini-CEX is, the higher the probability of achieving a superior performance is for each item.

Internal consistency

The total number of items in the instrument obtained a Cronbach's alpha of 0.74.

Acceptability

A statistically significant difference ($p < 0.001$) was found between the satisfaction level of supervisor assessors and peer assessors, having a median of 6 (6-6.75) and 7 (6-7), respectively. Student satisfaction was not affected when assessment by supervisors or by peers (Table 3).

Observation-time and feedback-time

A statistically significant difference was found in the observation-time, depending on whether the assessor was supervisors or peer. For supervisors, the median was 10 minutes (5-15), whereas for peers, the median was 30 minutes ($p < 0.001$). This difference is not observed when comparing the time needed for feedback (Table 3).

DISCUSSION

Students' educational outcomes and the healthcare they provide to patients could be improved by turning assessment into learning opportunities.³² In this respect, the Mini-CEX was adapted to the dentistry program's clerkship, and a psychometric analysis of its pilot implementation was carried out.

Some reports on the use of Mini-CEX in dentistry indicate that repeated implementation of it boosts students' confidence and that timely and individualized feedback improves clinical practice and patient care.^{6,25} Assessors consider that it allows assessing the capacity of reflection, self-perception,^{20,25} communication skills, and professionalism;¹² thus, conducting an assessment which is more objective than an only grade at the end of the semester.²³ The students evaluated the use of Mini-CEX positively.¹²

The existing literature on WBA recommends a longitudinal design with low-stake assessment for students to accumulate a sufficient number of encounters, which helps them know their weak areas and, consequently, improves their performance.^{20,33} Previous studies recommend to conduct between ten to fourteen application of Mini-CEX to obtain valid results. However, recent studies suggest that a lower number of assessments (in the range from four to ten) are considered valid for most assessment objectives.¹⁰ In this study, each student was assessed at least four times in fifteen weeks [median of 5 (4-5)].

For the current study, the research team decided to use a scale from 1 to 7, since a previous study in which the Mini-CEX was implemented in the physiotherapy clerkship concluded that this scale was the most culturally accepted in Chile.³⁴ This change did not affect the validity of the results, as the assessors' reliability is similar in scales of nine and five points.³⁵ Even though nine-point scales give more accurate scores than those of five points,³⁵ the literature indicates that validity can be improved by assigning labels to the different value ranges (insufficient, sufficient,

and superior).³⁶ This variation of the original nine-point scale of the Mini-CEX has been reported in other studies, where scales of 5, 7, 8 and 10 points-scale were used.¹⁰

The ceiling effect observed in the results, with a low dispersion of the scores (concentrating on the "superior" category in all the items), can be explained by a lack of training of the assessors (peers and supervisors). It has been concluded that inadequate training may impact assessors not giving scores in the lower level of the assessment form;^{1,12} which is called leniency-bias or generosity error.³⁷ An increase in the number of options on the scale has not been shown to solve this problem,³⁸ the literature recommends improving the training of assessors^{1,5,12,38} emphasizing the objectives of the assessment and the importance of the feedback to improve the student results and the care they provide to their patients.²⁹

In this study, a training session was carried for supervisors and peers, in which the observable behaviors were agreed to evaluate the different items. To improve the process of assessing clinical competencies in dental clerkship, a training plan focused on WBA and feedback delivery will be designed in the future.

The psychometric analysis of this adapted version of the Mini-CEX showed that it is a unidimensional instrument with an internal consistency of 0.74. These results are consistent with those reports in the literature which indicate that the Mini-CEX is unidimensional and its internal consistency varies from 0.59 to 0.97.¹⁰

With regard to the results of the observation-time (median of 15 minutes) and the time used in feedback (median of 5 minutes), they are similar to those reported in the literature, where the time of observation varies from 12.30 minutes to 46.5 minutes^{10,39,40} and the time used in feedback varies from 5.73 minutes and 20 minutes.^{10,25,39,40} It should be noted that in this study peer assessors had higher observation times than supervisors assessors, obtaining medians of 30 minutes and 10 minutes, respectively, this was considered a statistically significant difference.

Several studies reported experiences on the effectiveness of peer assessment in students of clinical courses of dentistry have shown that the students could properly identify in their peers the domains of superior performance and those in which improvement was needed,⁴¹ peer-assisted learning promotes a collaborative environment among students at similar or different levels of educations,^{42,43} engaging students as peer teachers could reinforce the learning experience for them.⁴⁴ It is essential to take into account the implications for student well-being and the effects on group cohesion when implementing peer evaluation of summative instances,^{1,45} institutional culture must emphasize security to assess and the student must be trained to provide adequate feedback.^{1,46}

A high level of acceptance from students and assessors has been reported in the literature.^{10,25} As to this study, high levels of acceptance from the students were also observed. Regarding the assessors, it was observed that there was a higher level of acceptance from peers than from supervisors; the level of acceptance was still high in both cases.

Additionally, it has been reported that there could be a percentage of incomplete data in paper forms.⁴⁷ In our study, the 8 domains of all forms were qualified; however, only a low percentage of the data on the observation time, feedback time and assessor and student satisfaction was completed.

The utility analysis for assessment methods involves the reliability, validity, educational impact, acceptability, and cost of the assessment tools.⁴⁸ The results of this research provide data on the validity, reliability, and acceptability of the Mini-CEX, in the context of the dental clerkship, which are similar to those reported for the original instrument.¹⁰ Considering the value of low-risk evaluations to make informed decisions about student performance, for future research, it is recommended to evaluate the educational impact to the dental clerk and predictive validity of the

Mini-CEX results, in addition to correlating the scores obtained with other assessment instruments and inter-rater reliability.

The present study has some limitations. Firstly, it was not possible to establish a correlation with other assessment instruments or to evaluate the inter-rater reliability. Secondly, it was not possible to assess the satisfaction level of students and assessors beyond the self-report of the form used. Thirdly, the participants did not complete all additional information (satisfaction, observation's and feedback's time), therefore the data could be biased. Lastly, for interpreting these results, it should be considered that the calculated sample size was not reached.

CONCLUSION

The results of this study suggest that the adapted version of the Mini-CEX, which was administered in dental interns, can objectively assess clinical performance, with values of validity and reliability similar to those of the original instrument. The collected data indicates that the adapted Mini-CEX operated as a unidimensional instrument with acceptable internal consistency and domains with a similar difficulty, in addition to having the acceptance of supervisors and students.

This Mini-CEX adaptation is a valid and reliable instrument for evaluating in dental clerkship; having a tool of this nature will optimize evaluation times and make the most clinical encounters to assess the interns' performance comprehensively and objectively.

ACKNOWLEDGMENTS

This research is part of a thesis leading to the Master of Medical Education and Health Sciences degree of the Faculty of Medicine at Pontificia Universidad Católica de Chile. The authors would like to thank the supervisors and students who participated in this study.

REFERENCES

1. Shumaway JM, Harden RM. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003;25(6):569–84.
2. Schuwirth LWT, Van Der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. *Med Teach*. 2011 Jun;33(6):478–85.
3. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012 Mar;34(3):205–14.
4. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(9):S63–7.
5. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*. 2007 Nov;29(9–10):855–71.
6. Kalsi HK, Kalsi JS, Fisher NL. An explanation of workplace-based assessments in postgraduate dental training and a review of the current literature. *Br Dent J*. 2013;215(10):519–24.
7. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): A preliminary investigation. *Ann Intern Med*. 1995;123(10):795–9.
8. Ansari A Al, Ali SK, Donnon T. The construct and criterion validity of the mini-CEX: A meta-analysis of the published research. *Acad Med*. 2013;88(3):413–20.
9. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The Mini-CEX: A Method for Assessing Clinical Skills. *Ann Intern Med*. 2003;138(6):476–81.
10. Mortaz Hejri S, Jalili M, Masoomi R, Shirazi M, et al. The utility of mini-Clinical Evaluation Exercise in undergraduate and postgraduate medical education: A BEME review: BEME Guide No. 59. *Med Teach*. 2019;0(0):1–18.
11. Cook DA, Beckman TJ, Mandrekar JN, Pankratz VS. Internal structure of mini-CEX

scores for internal medicine residents: Factor analysis and generalizability. *Adv Heal Sci Educ*. 2010;15(5):633–45.

12. Behere R. Introduction of mini-CEX in undergraduate dental education in India. *Educ Heal Chang Learn Pract*. 2014;27(3):262–8.

13. O'Connor A, McGarr O, Cantillon P, McCurtin A, et al. Clinical Performance Assessment Tools in Physiotherapy Practice Education: A Systematic Review. *Physiotherapy*. 2017 Mar 1;104(1):46–53.

14. Sheppard L, Crowe M, Jones A, Adams R. A model for re-education of physiotherapy practitioners to enable return to the workforce. *J Physiother*. 2010;30(1):23–31.

15. Wei CJ, Lu TH, Chien SC, Huang WT, et al. The development and use of a pharmacist-specific Mini-CEX for postgraduate year trainees in Taiwan. *BMC Med Educ*. 2019;19(1):1–8.

16. Liu YP, Jensen D, Chan CY, Wei CJ, et al. Development of a nursing-specific Mini-CEX and evaluation of the core competencies of new nurses in postgraduate year training programs in Taiwan. *BMC Med Educ*. 2019 Jul 18;19(1):270.

17. Milner KA, Watson SM, Stewart JG, DeNisco S. Use of Mini-CEX tool to assess clinical competence in family nurse practitioner students using undergraduate students as patients and doctoral students as evaluators. Vol. 53, *Journal of Nursing Education*. Slack Incorporated; 2014. p. 719–20.

18. Rogausch A, Beyeler C, Montagne S, Jucker-Kupper P, et al. The mini-clinical evaluation exercise during medical clerkships: Are learning needs and learning goals aligned? *Med Educ*. 2014;48(10):1008–19.

19. Rogausch A, Beyeler C, Montagne S, Jucker-Kupper P, et al. The influence of students' prior clinical skills and context characteristics on mini-CEX scores in clerkships - a multilevel analysis. *BMC Med Educ*. 2015;15(1):1–10.

20. Field JC, Walmsley AD, Paganelli C, McLoughlin J, et al. The Graduating European Dentist: Contemporaneous Methods of Teaching, Learning and Assessment in Dental Undergraduate Education. *Eur J Dent Educ*. 2017;21(July):28–35.
21. Mattheos N, Ucer C, Van De Velde T, Nattestad A. Assessment of knowledge and competencies related to implant dentistry in undergraduate and postgraduate university education. *Eur J Dent Educ*. 2009;13(SUPPL1.):56–65.
22. Prescott LE, Norcini JJ, Mckinlay P, Rennie JS. Facing the challenges of competency-based assessment of postgraduate dental training: Longitudinal Evaluation of Performance (LEP). *Med Educ*. 2002;36(6):92–7.
23. Iniesta M, Viñuela J, Utrilla F, Aracil L, et al. Implementation of Mini-Cex in the Subject of Periodontology. 5Th Int Conf Educ Res Innov (ICERI 2012). 2012;(November):4203–10.
24. Prescott-Clements L, van der Vleuten CPM, Schuwirth L, Gibb E, et al. Measuring the development of insight by dental health professionals in training using workplace-based assessment. *Eur J Dent Educ*. 2011;15:159–64.
25. Gerhard-Szep S, Güntsch A, Pospiech P, Söhnel A, et al. Assessment formats in dental medicine: An overview. *GMS Z Med Ausbild*. 2016;33(4):1–43.
26. Vázquez M, García L, Pinchak C, Peluffo G, et al. Evaluación de las competencias clínicas de estudiantes de Posgrado de Pediatría utilizando el Mini-Clinical Evaluation Exercise (Mini-CEX). *Arch Argent Pediatr*. 2011 Mar 5;109(4):314–20.
27. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods*. 1996 Jan 1;1(2):130–49.
28. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000 Oct 1;32(4):1008–15.
29. Reise SP, Waller NG. Item Response Theory and Clinical Measurement. *Annu Rev Clin*

Psychol. 2009 Apr;5(1):27–48.

30. Artino AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach*. 2014;36(6):463–74.

31. Nunnally J., Bernstein I. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill, Inc. 3rd ed. McGraw-Hill, editor. New York; 1994.

32. Harden RM, Grant J, Buckley G, Hart IR. BEME Guide No. 1: Best evidence medical education. Vol. 21, *Medical Teacher*. *Med Teach*; 1999. p. 553–62.

33. Albino JEN., Young S, Neumann LM, Kramer GA, et al. Assessing Dental Students' Competence: Best Practice Recommendations in the Performance Assessment Literature and Investigation of Current Practices in Predoctoral Dental Education. *J Dent Educ*. 2008;77(12):1405–35.

34. Fuentes J. Evaluación para el aprendizaje: diseño e implementación de un Mini-CEX en el internado profesional de la Carrera de Kinesiología de la Pontificia Universidad Católica de Chile. [Magister]: Pontificia Universidad Católica de Chile; 2017.

35. Cook DA, Beckman TJ. Does scale length matter? A comparison of nine- versus five-point rating scales for the mini-CEX. *Adv Heal Sci Educ*. 2009 Dec 26;14(5):655–64.

36. Hancock GR, Klockars AJ. The effect of scale manipulations on validity: Targeting frequency rating scales for anticipated performance levels. *Appl Ergon*. 1991 Jun 1;22(3):147–54.

37. Pangaro L, Ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. *Med Teach*. 2013;35(6).

38. Keeley JW, English T, Irons J, Henslee AM. Investigating Halo and Ceiling Effects in Student Evaluations of Instruction. *Educ Psychol Meas*. 2013 Jun 4;73(3):440–57.

39. Al-Jewair T, Kumar S. Review and Application of the Mini-Clinical Evaluation Exercise

(Mini-CEX) in Advanced Orthodontic Education: A Pilot Study. *J Dent Educ.* 2019 Jul

22;83(11):1332–8.

40. Deshpande S, Pande N, Radke U, Jaiswal P, et al. Evaluation of Mini Clinical Evaluation Exercise (Mini-CEX) for post-graduate students in Prosthodontics: Student and Faculty Perceptions. *MedEdPublish.* 2018;7(1):1–8.

41. Tricio J, Woolford M, Thomas M, Georghiou L, et al. Dental students ' peer assessment : a prospective pilot study. *Eur J Dent Educ.* 2015;19:140–8.

42. Walker-Bartnick LA, Berger JH, Kappelman MM. A model for peer tutoring in the medical school setting. *J Med Educ.* 1984;59(4):309–15.

43. Buckley S, Coleman J, Davison I, Khan KS, et al. The educational effects of portfolios on undergraduate student learning: A Best Evidence Medical Education (BEME) systematic review. BEME Guide No. 11. *Med Teach.* 2009;31(4):282–98.

44. Kim AH, Chutinan S, Park SE. Assessment Skills of Dental Students as Peer Evaluators. *J Dent Educ.* 2015 Jun 1;79(6):653–7.

45. Stagg K. Impact of peer assessment on student wellbeing [Internet]. Vol. 38, *Medical Teacher.* Taylor and Francis Ltd; 2016. p. 212.

46. Nofziger AC, Naumburg EH, Davis BJ, Mooney CJ, et al. Impact of peer assessment on the professional development of medical students: A qualitative study. *Acad Med.* 2010;85(1):140–7.

47. Chang YC, Chen CK, Chen JC, Liao CH, et al. Implementation of the mini-clinical evaluation exercise in postgraduate Year1 residency training in emergency medicine: Clinical experience at Chang Gung Memorial Hospital. *J Acute Med.* 2013 Sep 1;3(3):110–5.

48. Van Der Vleuten CPM. The assessment of professional competence: Developments, research and practical implications. *Adv Heal Sci Educ.* 1996;1(1):41–67.

Table 1. Descriptor of domains and observable behaviors of the student

Domain	Description	Behavior observed by the assessor.
Anamnesis and physical examination focused on the patient's condition.	Makes an anamnesis that is efficient and relevant to the patient's health condition, following a logic sequence (good time management).	<ul style="list-style-type: none"> -The questions asked to the patient are coherent with the information needed to elaborate a correct diagnosis. No missing questions or irrelevant information is asked for. -The time spent for the patient's anamnesis is coherent with the patient's condition.
Clinical judgement	Explains the treatment planning, using their knowledge for therapeutic decision-making.	<ul style="list-style-type: none"> -Responds correctly and with confidence the questions about the reasons and indication of the treatment that are asked by the supervisors. -Bases their actions on the evidence available for the case.
Autonomous execution of the treatment according to level	Executes the planned treatment autonomously according to level, complying with the established protocols depending on the specific diagnosis, and biosafety regulations (workspace and patient and personal protective equipment).	<ul style="list-style-type: none"> -When executing the treatment, the student asks the supervisors for guidance only if needed, considering the complexity of the action. -Complies with the protocols defined for the different treatments. -Complies with all biosafety regulations (patient and personal protection) -Keeps the order in each of the box area (dirty area, clean area and workspace)
Planning and organization of the session	<ul style="list-style-type: none"> Prioritizes properly the planned actions by order of importance. Organizes efficiently the session and accomplishes the planning. 	<ul style="list-style-type: none"> -When planning the actions that will be performed in a clinical session, the student prioritizes them properly according to the action importance, patient's preferences and global treatment planning. - When planning the actions that will be performed in a clinical session, the student considers the time needed for each action correctly (neither time spared nor lack of time)

Efficient use of resources	Identifies and uses the resources efficiently (materials and supplies).	<p>-When asking for supplies in the first aid kit, the student does it efficiently. The students do not ask for supplies in excess (type or quantity)</p> <p>-Identifies and handles needed biomaterials correctly and uses the equipment in a due manner to accomplish the planning.</p>
Professionalism	<p>Shows respect and empathy and attends the patient's needs in terms of wellbeing and information.</p> <p>Acknowledges their limitations demonstrates self-criticism and asks for help when it is necessary.</p>	<p>-When interacting with the patient, the student can show respect and empathy, explains the procedures that will be conducted in proper terms (keeps eye-contact during the interaction), does not leave the patient alone and unattended gives the information needed to understand the order of the treatment and indicates the current stage</p> <p>-Addresses the patient by their name.</p> <p>-When assessing its own work, the student demonstrates self-criticism (based on the quality of the work performed)).</p> <p>-Asks for help when the complexity of the action demands the guidance of their supervisors.</p>
Communication with the work team and patient	Communicates with their peer and supervisors. Responds properly to verbal and non-verbal signals or cues from the patient and is capable of dealing with difficult situations and solve them (negotiation, treatment plan, expectation management) if any.	<p>-Communicates with their peer and supervisors respectfully, using vocabulary and tone of voice according to the situation.</p> <p>-The student can negotiate with the patient the order of the treatment, weighing the importance of the action and the patient's preferences properly.</p> <p>-The student can properly read the non-verbal signals from the patient when the patient expresses fear, anxiety or doubt, among others, and responds to them in due manner, keeping the patient calm.</p>
Global competence	Informs the general perception of the observer on student performance.	

Table 2. Standardized factor loadings and R^2 associated to the instrument items.

Domain	Standardized factor loading (exploratory analysis)	Standardized factor loading (confirmatory analysis)	R^2
1.-Anamnesis and physical examination focused on the patient's condition.	not analyzed ^a	not analyzed ^a	-
2.-Clinical judgement	0.93	0.93	0.87
3.-Autonomous execution of the treatment according to level.	0.84	0.87	0.79
4.-Planning and organization of the session.	0.43	0.36	0.13
5.-Efficient use of resources.	-0.03	0.03	0.001
6.-Professionalism.	0.34	0.30	0.09
7.-Communication with the work team and patient.	0.01	0.02	0.0003
8.-Global competence.	0.89	0.88	0.78

^a This item was excluded since there exists no variability.

Table 3 Satisfaction, observation time and feedback time

Item	Total			Instruments applied by supervisors			Instruments applied by peers		
	Median	Q1 ^a	Q3 ^b	Median	Q1 ^a	Q3 ^b	Median	Q1 ^a	Q3 ^b
Assessor satisfaction with the instrument	6	6	7	6*	6	6.75	7*	6	7
Student satisfaction with the instrument	7	6	7	7	6	7	7	7	7
Observation time	15	10	30	10*	5	15	30*	20	45
Time needed for feedback	5	5	7	5	5	7	5.5	5	9.25

*Statistically significant differences between peers and supervisors $p < 0.001$

^a First quintile

^b Third quintile

Student: _____ Date: _____

Assessor: _____ Tutor ☐ Peer ☐

Diagnosis ☐ Treatment ☐ Check-up ☐ Performed activity: _____

1.- Anamnesis and/or physical examination of the patient's condition

1	2	3	4	5	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

2.- Clinical judgement

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

3.- Autonomous execution of the treatment according to level

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

4.- Planning and organization of the session

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

5.-Efficient use of resources

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

6.-Professionalism

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

7.-Communication with the work team and patient

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

8.-Global competence

1	2	3	4	0	6	7	0
Unsatisfactory			Satisfactory		Superior		Not observed

Application time ☐ Observation time ☐ Feedback time ☐

Assessor satisfaction with the instrument **Student satisfaction with the instrument**

Low 1 2 3 4 5 6 7 High Low 1 2 3 4 5 6 7 High

Comments/Action plan

Assessor Signature _____ Student Signature _____

Figure 1. Adapted instrument

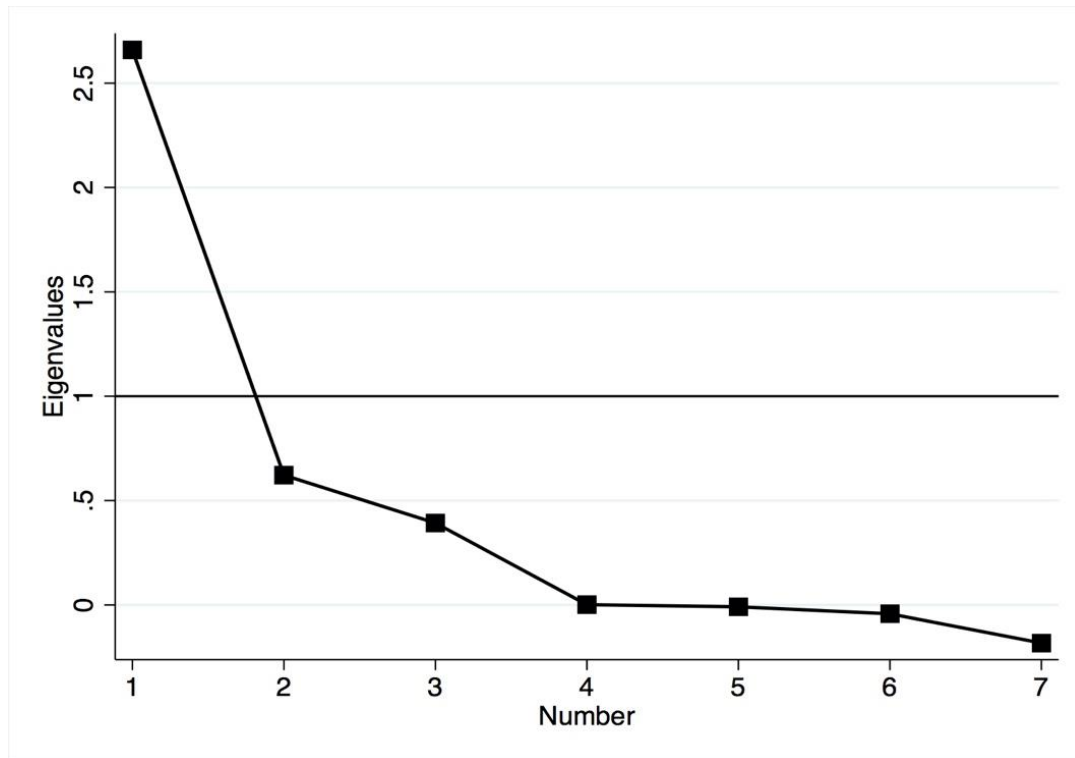


Figure 2. Sediment graph that shows Eigenvalues (axis y) and the number of factors (axis x) of the scale.

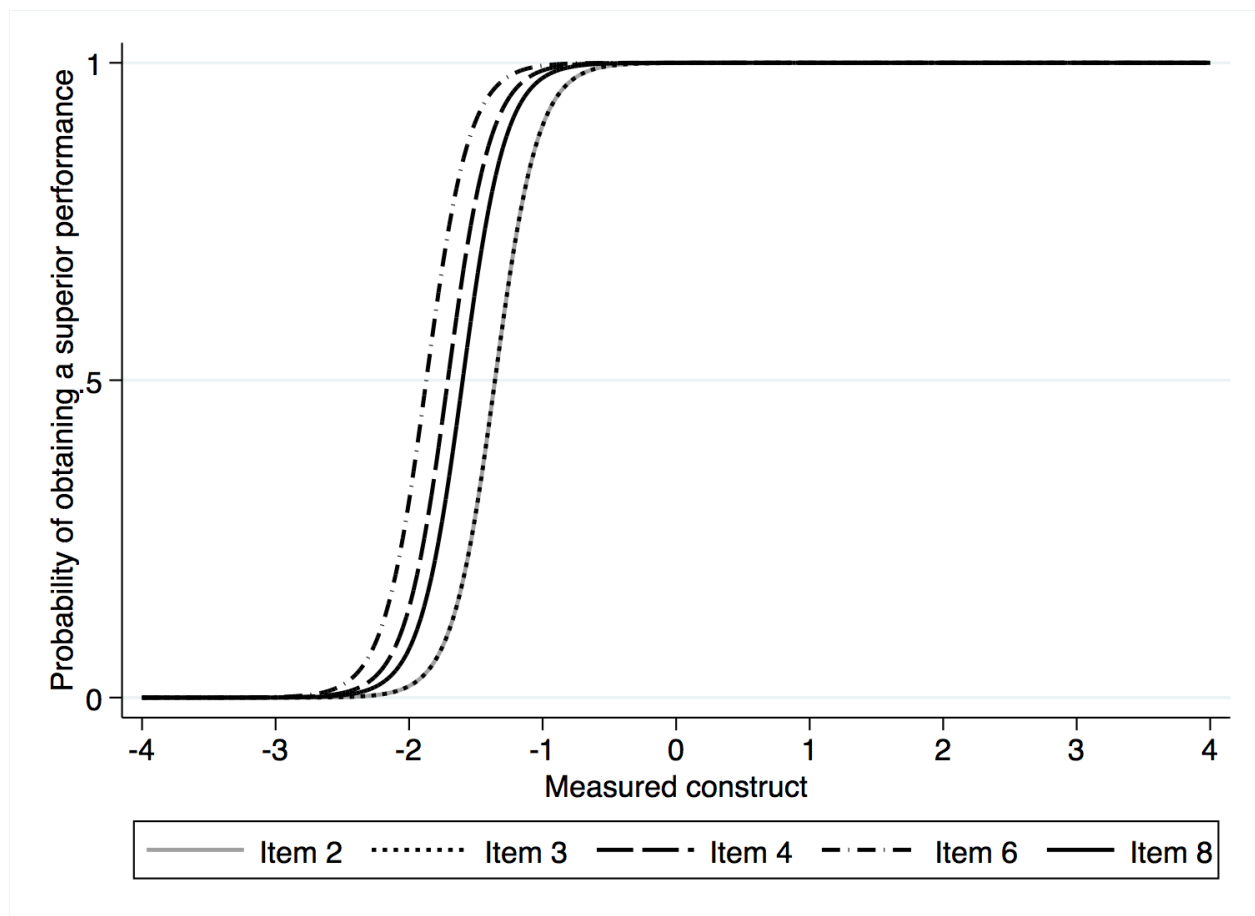


Figure 3. Item characteristics curves for the 5 items with variability of scores.