Evaluation of Clinical and Surgical Proficiency at a Residency of Child Orthopedics and Traumatology

Use of the Mini-CEX (*Mini-Clinical Evaluation Exercise*) and DOPS (*Direct Observation of Procedural Skills*)

Claudio A. Fernández, María Gabriela Miranda
Orthopedics and Traumatology Service, Children's Hospital "Sor María Ludovica" (La Plata, Buenos Aires, Argentina)

ABSTRACT

Purpose: To assess professional competencies in a Pediatric Orthopedic and Traumatology Residency Program by the implementation of two performance-assessment instruments: Mini-Clinical Evaluation Exercise (Mini-CEX) and Direct Observation of Procedural Skills (DOPS) for clinical and surgical skills. Both tools aim to assess the top of Miller's pyramid, for its reliability and validity. **Materials and Methods:** Prospective observational cohort study of six medical trainees in the first, second and third year of their residency program (R1-R2-R3) who were randomly assessed by six examiners during their daily training at outpatient clinics, emergency room, inpatients unit, operating room, and plaster room. The statistical analysis was carried out with the Chi-Square and Wilcoxon-Rank paired test for univariate variables. The residents' relationship cohorts were compared using the Kruskal-Wallis test. The reliability of the methodological tool was determined by the psychometric test of Crombach. Alfa was set at ≤ 0.05. Diagnostic study: level IV. **Results:** We performed 65 assessments. Each resident was evaluated 10 times on average by 3 to 6 examiners. The oldest residents had better performances in overall clinical competencies. However, the R1 group achieved satisfactory results whereas the R2-R3 groups had the most outstanding scores. There were no statistical differences in general surgical competencies, but the R3 group was outstanding in cases of unforeseen surgical situations. The Alfa Crombach coefficient was over 0.90. **Conclusion:** The Mini-CEX, DOPS, and interactive feedback were powerful tools to provide high-quality assessment and were widely accepted by residents and examiners. The statistical analysis allowed us to identify the weaknesses and strengths of the trainees. The Crombach coefficient had a high psychometric impact.

Keywords: Residency; assessment; competencies; Mini-CEX; DOPS.
Level of Evidence: IV

Evaluación de competencias clínicas y quirúrgicas de una Residencia de Ortopedia y Traumatología Infantil. Utilización del Mini-CEX (*Mini-Clinical Evaluation Exercise*) y del DOPS (*Direct Observation of Procedural Skills*)

RESUMEN

Objetivo: Evaluar competencias profesionales de una residencia de Ortopedia y Traumatología Infantil. Instrumentos pedagógicos utilizados: *Mini-Clinical Evaluation Exercise* (Mini-CEX) y *Direct Observation of Procedural Skills* (DOPS) para competencias clínicas y quirúrgicas, respectivamente. Ambas evalúan la cúspide de la pirámide de Miller; se precisaron su confiabilidad y validez. **Materiales y Métodos:** Estudio observacional prospectivo de una cohorte de seis residentes de primero, segundo y tercer año quienes fueron evaluados por seis docentes en diversos contextos y situaciones reales: consultorio externo y de guardia, sala de internación, interconsultas, quirófano y sala de yesos. **Resultados:** Se realizaron 65 observaciones. Cada residente fue evaluado como media en 10 oportunidades por entre 3 y 6 docentes. Para las variables clínicas, en general, los residentes más antiguos obtuvieron valores sobresalientes y los residentes de primer año, valores satisfactorios. No hubo diferencias significativas para las competencias quirúrgicas globalmente, pero los residentes de tercer año fueron más competentes para resolver situa-

Received on August 20th, 2020. Accepted after evaluation on October 26th, 2020 • CLAUDIO A. FERNÁNDEZ, MD • claufernandez@outlook.com https://orcid.org/0000-0003-2350-3885

How to cite this article: Fernández CA, Miranda MG. Evaluation of Clinical and Surgical Proficiency at a Residency of Child Orthopedics and Traumatology. Use of the Mini-CEX *(Mini-Clinical Evaluation Exercise* and DOPS *(Direct Observation of Procedural Skills). Rev Asoc Argent Ortop Traumatol* 2021;86(3):417-427. https://doi.org/10.15417/issn.1852-7434.2021.86.3.1188

ciones inesperadas. El coeficiente de Cronbach fue superior a 0,90. **Conclusiones:** Ambos instrumentos de evaluación tuvieron una elevada confiabilidad. El método estadístico permitió individualizar exactamente las fragilidades y fortalezas de la residencia. El coeficiente de Cronbach obtuvo un valor de alto impacto psicométrico.

Palabras clave: Residencia; evaluación; competencias; Mini-CEX; DOPS.
Nivel de Evidencia: IV

INTRODUCTION

Medical education in Child Orthopedics and Traumatology consists of the development and acquisition of specific competencies, the purpose of any curriculum planned and included in the official program of a residency. Competence is defined as the set of knowledge, skills and attitudes that provide excellent medical practice in continuous improvement and appropriate to the social context in which it takes place. It includes cognitive aspects (knowing and understanding), psychomotor skills (knowing how to act), and training in values and attitudes (knowing how to be).[1,2]

Determining whether resident physicians meet the expectations of achievement requires the use of structured, adaptable and relevant assessment tools in relation to the content and objectives included in the training curricula. The fundamental purpose of an assessment is to analyze the learning process, something we do *with* and *for* the residents—not *to* the residents.[1,3,4]

In 2010, the Consensus Group at the Ottawa Conference defined the criteria for proper evaluation: validity, reliability, fairness and equity, equivalence, feasibility, educational and catalytic effect, and acceptability. A tool is valid when it effectively evaluates what it intends to evaluate, and is consistent with teaching strategies and content. Reliability is a statistical concept that defines the stability or reproducibility of a test; it is expressed by a correlation coefficient, in which 1 is perfection and 0 is nullity. Justice and equity respond to the social principle of education. Equivalence is the probability of application of the same assessment methodology in subsequent meetings or different institutions. Feasibility is the availability of physical, human and financial resources. Every evaluation affects the educational strategy and the future of the professional. The catalytic effect is the ability to imitate or spread an assessment methodology.[5]

In 1990, Miller presented a four-step pyramid of assessment (Figure). The lower two tiers are formed by theoretical knowledge *(knows)* and its application in specific cases *(knows how).* The upper tiers refer to behavior: the third expresses competence in simulated or *in vitro* environments (*shows how*), whereas the fourth tier or pinnacle of the pyramid is the reference pattern of professionalism or the practice of medicine in real situations *(does).* It is feasible to evaluate each item by specific instruments. The framework for the assessment at the pinnacle of the pyramid includes: *Clinical Evaluation Exercise* (CEX), *Mini-Clinical Evaluation Exercise* (Mini-CEX), *Direct Observation of Procedural Skills* (DOPS), feedback or *multi-source feedback* 360, portfolio, case-based studies, video recordings, peer evaluations, audits and others. All can be used for formative or summative assessment, or both (Figure).[6,7]

The direct observation of the clinical and surgical skills of resident physicians is essential. To this end the *American Board of Internal Medicine* developed the Mini-CEX in 1995.[8-11] This is a structured assessment tool that involves patients and real situations in varied contexts and with diverse difficulties. The DOPS tool was introduced in 2005, in the United Kingdom, by the *United Kingdom Foundation Programme*, as a resource for assessing specific surgical and technical competencies. It also uses patients and real-world situations in scenarios of varying complexity.[12-18] It is important to note that it is not designed to assess *in vitro* situations in animals or in anatomical theatre.[9] The Mini-CEX has received more support from the literature, probably because the publication and application of DOPS are newer.[4,8]

These assessments can be conducted at the request of the observer or resident, which significantly reduces psychic stress or anxiety from the experience.[8] Both include a systematic return or feedback.[8,9,16]

The purpose of this research was to determine the validity and reliability of the Mini-CEX and DOPS tools in *a training assessment* of specific clinical, surgical and orthopedic practices at a residency of Child Orthopedics and Traumatology.

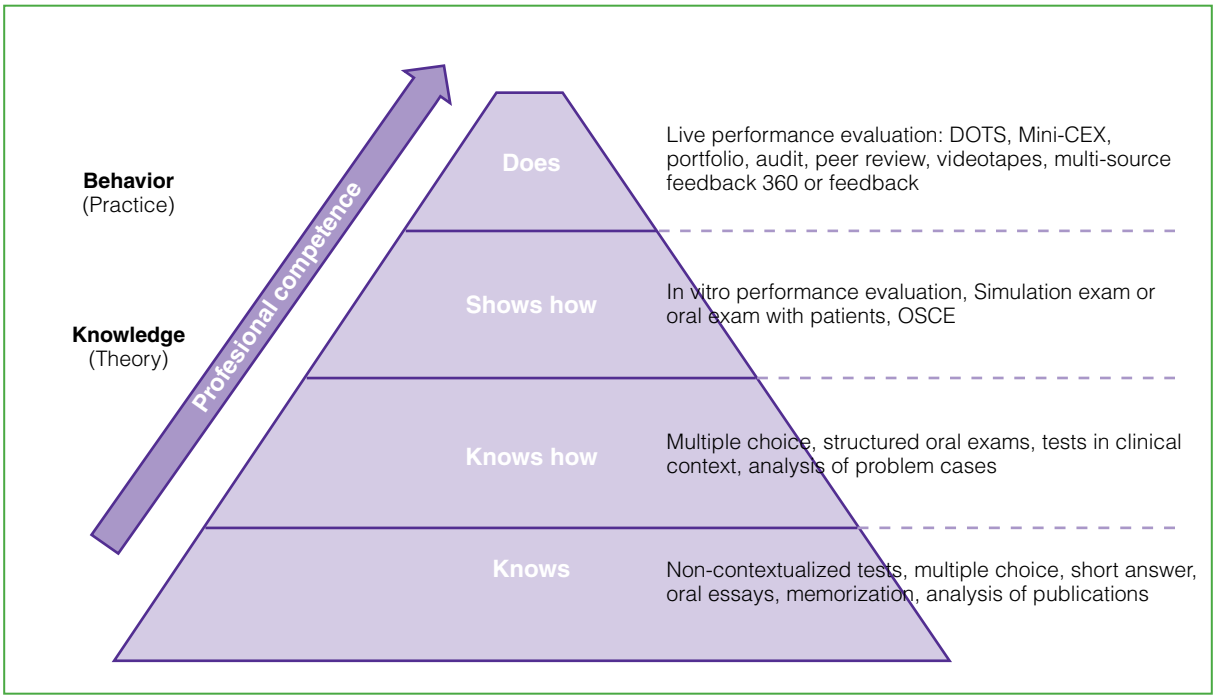

Figure. Miller's pyramid of knowledge and correlation with evaluation tools.

MATERIALS AND METHODS

It is an observational and prospective cohort study, of the diagnostic category. Between November 2018 and April 2019, doctors from a residency of Child Orthopedics and Traumatology—accredited by the *Asociación Argentina de Ortopedia y Traumatología* and the *Sociedad Argentina de Ortopedia y Traumatología Infantil*—were evaluated. It is a second-stage residency, the *sine qua non* requirement of which is to have undergone and passed a general Orthopedics and Traumatology residency, as well as a multiple-choice question examination. It is the only education system developed in a purely pediatric hospital of maximum complexity in the province. The program lasts for two years, with the option of a third year as head of residents. For the data collection and statistical processing, we will refer to first- and second-year resident physicians, and heads of residents, as R1, R2 and R3, respectively.

The body of evaluators consisted of staff physicians, who randomly chose both the resident to interview and the competence to assess. Conversely, some residents requested to be evaluated at a given activity. Mini-CEX was used for clinical competencies in various contexts: external and on-call practice, interconsultations and inpatient rooms. The complexity of the procedure was classified as low, medium or high. The following variables were included: anamnesis and physical examination, ethics or professionalism, clinical judgment, organization and efficiency. Each competence was valued according to the following scale: 1 to 3 or unsatisfactory, 4 to 6 or satisfactory and 7 to 9 or outstanding. The time of observation and feedback was allocated. Both evaluator and resident were able to express their satisfaction with the encounter on a scale of 1 to 10. The observer was able to point out in writing the positive aspects of the experience and those that could be improved (Table 1).

Table 1. *Mini-Clinical Evaluation Exercise* (Mini-Cex) structured form.

Mini-Cex (Mini Clinical examination) Structured observation of clinical practice											
Evaluator:					Date						
Resident					R1 ( )		R2 ( )		R3 ( )		
Environment	Outpatient office ( )			Emergency ( )			Room ( )			Referral ( )	
Patient	First time ( )			Follow-up ( )			Age			Gender	
Diagnosis:											
Complexity	Low ( )		Medium ( )			High ( )					
Observer	Staff Physician ( )			Professor ( )							
1. **Anamnesis**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
2. **Physical examination**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
3. **Profesionalism**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
4. **Clinical judgement**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
5. **Communication skills**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
6. **Organization/ Efficiency**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
7. **Global score**										**Not evaluated**	**Observation**
1	2	3		4	5	6		7	8	9	
Unsatisfactory				Satisfactory				Superior			
• Particularly positive aspects:											
• Aspects that need to be improved:											
• Resident's satisfaction with the MiniCex 1 2 3 4 5 6 7 8 9 10											
• Observer's satisfaction with the MiniCex 1 2 3 4 5 6 7 8 9 10											
• Time (minutes) used for observation:											
• Time (minutes) used for feedback:											
• Comments:											
• Date:											
• Resident's signature:									• Observer's signature:		

DOPS was chosen to evaluate competencies in scheduled and on-call surgeries, as well as in the orthopedic treatment of clubfoot with the Ponseti technique. The residents were assigned a role as surgeon or assistant. The complexity of the procedures was classified into: less than usual, usual and more than usual. The variables evaluated were: specific informed consent, therapeutic indication, anatomy, procedure and complications, preparation of the procedure according to protocol, demonstration of good asepsis and proper use of the instruments, orderly development of the technique, respect for anatomical structures, behavior in an unexpected situation, communication with the surgical team, preparation of surgical protocol, demonstration of professional behavior in all instances and indications of post-surgical care. Each observation was awarded the following score: 0 = Not observed, 1 = lower-than-expected development, 2 = satisfactory or desired level and 3 = outstanding. Just as in the Mini-CEX, immediate feedback was given; the evaluator and the resident were able to express their satisfaction with the test, and comment on aspects to be considered and modified (Table 2). For both situations, the patient's diagnosis, age and sex were needed.

Table 2. *Direct Observation of Procedural Skills* (DOPS) structured form.

Direct observation of procedural skills. (Surgery and bot foot) (**DOPS**: Direct Observation of Procedural Skills)
Evaluator: Date
Resident: R1 ( ) R2 ( ) R3 ( )
Environment: Operating room ( ) Bot foot Clinic ( ) Plaster room ( )
Surgeon ( ) Assistant ( )
Patient First time ( ) Follow-up ( ) Age Gender
Diagnosis:
Complexity of the procedure: Less than usual ( ) Usual ( ) More than usual ( )
Name of the procedure:
Elective surgery ( ) Emergency ( ) Closed reduction of fractures / dislocations ( )
Observer Staff Physician ( ) Professor ( ) Other ( )

0 = Not observed, 1 = Lower-than-expected development, 2 = Satisfactory or desired level 3 = Outstanding

Competencies	**Score**	**Comment**
1. Obtains informed consent and lists possible complications		
2. Describes the indication, anatomy, procedure and complications		
3. Prepares the procedure according to protocols		
4. Demonstrates good asepsis and uses instruments safely		
5. Develops the technique in an orderly manner, respects the anatomical structures		
6. Behaves appropriately in unexpected situations		
7. Communicates clearly with the surgical team and staff		
8. Fulfills the surgical protocol		
9. Behaves professionally at all times		
10. Indicates post-surgical care		

Comments
• Particularly positive aspects:
• Aspects that need to be improved:
• Recommendations:
• Resident's satisfaction with the DOPS 1 2 3 4 5 6 7 8 9 10
• Observer's satisfaction with the DOPS 1 2 3 4 5 6 7 8 9 10
• Time (minutes) used for observation:
• Time (minutes) used for feedback:
• Comments:
• Date: .....
• Resident's signature: ..... • Observer's signature: .....

An explanation of the tests to be used, their objectives and importance, as well as the structured files, were sent in advance by email, both to residents and the body of evaluators. Subsequently, a face-to-face explanatory meeting was held with evaluators and residents, in order to train them with the methodology and resolve doubts, emphasizing the concept that an evaluation is not an exam, it does not grant grades or certifications.

The data collected were transferred to Microsoft Windows 10 Excel spreadsheets for further analysis. The SPSS 17 program was used for statistical processing.[17] A significance level of $p \leq 0.05$ was established. The chi-square test sensitized to the Wilcoxon rank sum test for paired samples was used for the univariate analysis. The comparative analysis between groups of residents was performed with the Kruskal-Wallis test. The reliability of the assessment tools was determined by Cronbach's coefficient α.

FINDINGS

65 observations were made in six residents, by six staff physicians. Four of them are professors of a National University and all of them completed a medical residency (or equivalent) in Child Orthopedics and Traumatology. The average age of graduation was 25 years (range 13-38).

Mini-CEX: comprised 40 evaluations of 40 patients with an average age of 7.2 years (range6 months-14 years), without predominance of any sex. The observations corresponded to: 47.5% to R1, 22.5% to R2 and 30% to R3. Each evaluator made an average of 6.66 observations (range 5-10). The average time of the encounter was 16.4 min (range 10-30) and the time spent for feedback, 7 min (range 5-15). Each resident underwent an average of 6.66 assessments (range 3-10), in charge of 3 to 6 evaluators. Table 3 shows the number of observations made in each clinical scenario. The complexity of the consultation was: low (37.5%), moderate (32.5%), high (30%). 22.5% of consultations were "first-time"; 35.7% were "follow-up" and 42.5% were not recorded. Of all the competencies evaluated, the maximum score corresponded to professionalism and the lowest was related to clinical judgment and effective organization. R1s scored lower on the physical examination; R2s, in professionalism and R3s, in anamnesis (Table 4). However, there was no statistically significant difference in the observations of the residency as a whole. Evaluator satisfaction averaged 8.37 (range 4-10) and resident satisfaction averaged 8 (range 2-10) ($p < 0.003$).

Table 3. Mini-CEX. Frequency of evaluations for each clinical scenario

Scope of observation (n = 40)	Frequency	%
Outpatient office	15	37.5
Emergency	3	7.5
Room	16	40
Referral	6	15

Table 4. Scores by year of residence and scope of evaluation

Scope	1st year resident	2nd year resident	3rd year resident	p
Room	42	104	240	0.023
Referral	74	84	105	0.031
Outpatient office	502	349	217	0.027
Emergency	40	49	51	0.5

Inferential statistics: when specific competencies were analyzed by year of residency, R1s had lower overall scores, except in professionalism. R2s had lower scores in professionalism and R3s had slightly lower scores in anamnesis, although the difference between groups was not significant ($p = 0.31$). In assessing all competencies between groups, the R3s obtained significantly higher scores over the other groups ($p < 0.05$). Still, the R1s achieved a satisfactory average score; whereas the R2s and R3s were outstanding. Compared according to the field of observation, R1 and R2 obtained better scores in the outpatient clinic ($p < 0.02$) and R3, in the hospitalization room and interconsultations ($p < 0.03$). Cronbach's coefficient α was 0.92 for the Mini-CEX and 0.85 to 0.97 for the intrinsic relationship between variables (Tables 5 and 6).

Table 5. Mini-CEX. Competencies evaluated by scope, average values ($\bar{x}$)

Variables	Outpatient office	Emergency	Room	Referral	p
	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$	
Global score	54.9	49	48.25	37.57	0.5
Observation time (min)	16.5 (5-22)	15 (5-30)	16.4 (15-20)	16.25 (5-20)	0.4
Time used for feedback (min)	6.8 (5-10)	5 (3-8)	14.5 (10-30)	10 (5-15)	0.5
Resident's satisfaction	8.5 (6-10)	7.6 (6-10)	7 (4-10)	8.8 (5-10)	0.03
Evaluator's satisfaction	8.3 (4-10)	9 (6-10)	7.5 (4-10)	9.8 (7-10)	0.03

Table 6. Mini-CEX. Competences by year of residency

Evaluated area		R1		R2		R3		
Competencies	**n**	**Average**	**n**	**Average**	**n**	**Average**	**Average**	**p**
Anamnesis	95	5	80	7.2	70	7	19.2	0.3
Physical examination	92	4.8	76	6.9	80	8	19.7	0.3
Professionalism	119	6.2	83	4.5	79	7.9	18.6	0.3
Clinical judgement	114	6	89	8.09	78	7.8	21.8	0.3
Cognitive abilities	109	5.7	83	7.5	81	8.1	21.3	0.3
Effective organization	112	5.8	65	5.9	80	8	19.7	0.3
Global score	116	6.10	79	7.9	70	7	21	0.05

DOPS: over 25 observations, 40% corresponded to R1; 28% to R2; and 32% to R3. The evaluation included 23 patients, 78% were male, and the average age was 6.76 years (range 6 months-14 years). The average time of evaluation was 1 h and 21 min (range 10 min-5 h 50 min), the average feedback took 8.8 min (range 5-15). The complexity of the procedure was considered usual at 68% and more than usual at 32%. Residents performed as surgeons in 72% of the procedures and as assistants in the remaining procedures. On 19 occasions, trauma and orthopedic surgeries of various aetiologies were evaluated. In six cases, the evaluation included the treatment of bot foot with Ponseti technique, including manipulations, plasters and, in some patients, Achilles tendon tenotomy, as well as expertise with the use of the abduction splint. Each resident was evaluated, on average, 4.16 times (range 3-6) by 2.8 evaluators (range 2-4). Overall, 72% had a satisfactory score, and 28% did so outstandingly. None of them had a lower-than-expected result. Evaluator satisfaction with observation averaged 8 (range 7-10) and resident satisfaction averaged of 8.63 (range 6-10) ($p < 0.03$).

Inferential statistics: comparing resident groups, we see a statistically significant difference in overall competencies in favor of R3 ($p < 0.05$). Similarly, the latter group was more competent in resolving unexpected situations during the surgical act ($p < 0.005$). Cronbach α coefficient was 0.90 for DOPS, but for internal correlation between variables fluctuated between 0.89 and 0.93 (Tables 7 and 8).

Table 7. DOPS. Relationship between competencies in the surgical setting

Competencies	R1	R2	R3	Average	p
Global score	16.8	17	21	18.2	0.05
Number of observations	9	8	7	4.16	0.5
Average observation time (min)	66.1	60.8	49.2	58.7	0.5
Time used for feedback (min)	13 (5-15)	10 (5-15)	10.71 (5-15)	8.8	0.5
Resident's satisfaction	8	9	9	8.6	0.03
Evaluator's satisfaction	9	9	9	9	0.03

Table 8. Relationship between the years of residency and the surgical competencies analyzed.

Competencies	Scores (n)	R1	R2	R3	p
1. Obtains informed consent and lists possible complications	58	13	18	27	0.01
2. Describes the indication, anatomy, procedure and complications	47	13	16	18	0.5
3. Prepares the procedure according to protocols	56	10	16	30	0.01
4. Demonstrates good asepsis and uses instruments safely	65	17	18	30	0.01
5. Develops the technique in an orderly manner, respects the anatomical structures	68	18	20	30	0.02
6. Behaves appropriately in unexpected situations	68	18	20	30	0.005
7. Communicates clearly with the surgical team and staff	60	18	20	22	0.5
8. Fulfills the surgical protocol	65	18	20	27	0.5
9. Behaves professionally at all times	60	10	20	30	0.01
10. Indicates post-surgical care	68	18	20	30	0.05

DISCUSSION

The purpose of a residency program is the professionalization of the physician through the progressive acquisition of specific competencies. It is imperative to determine whether teaching strategies have been adequate and expectations of achievement have been met. This requires a process of systematic evaluation, diagnosis or data collection, by means of the application of various tools. Depending on the case, the evaluation can be diagnostic or initial, summative or final and formative, when it is desired to monitor the process and progress of learning. In our specialty, the task is arduous, because we must include very heterogeneous competencies. In *Civilization and Its Discontents,* Sigmund Freud ironically defined as impossible professions those in which the results are never completely satisfactory: psychoanalysis, government and education.[18]

We have already proposed that every tier of Miller's pyramid should be evaluated with relevant or specific instruments. Seeking to evaluate professionalism or even more, suitability, by means of a multiple-choice exam or written assignments, which refer exclusively to theoretical knowledge or, at most, to the "know-how", is irrelevant. They say nothing about "demonstrating" or "doing." The risk is evident. With theoretical knowledge, memory, correct answer recognition and prior practice, an individual can obtain excellent qualifications, even if this does not guarantee the quality of his/her professional practice at all. This is a false positive, which gives

professional credit to those who do not deserve it. We usually call this *Funes the Memorious Complex*, after the famous Jorge Luis Borges short story, in which he tells us the vicissitudes of Ireneo Funes, a young Uruguayan from Fray Bentos who, crippled as a result of an equestrian accident, develops hypermnesia or extraordinary memory, although lacking rational competencies. Insightful, Borges writes: "I suspect, however, that he was not very capable of thinking. To think is to forget differences, it is to generalize, to abstract."[19]

The practice of medicine is closely related to philosophy. Thus, when we reason we apply logic; by considering the patient as a real subject and external to our consciousness, we invoke naive realism; when we accept the existence of genes, microbes, technology and surgical techniques, we employ scientific realism; by rejecting the hypothesis of the mythical origin of diseases, we subscribe to naturalistic principles and, when we help beyond our own interests, we practice humanistic moral philosophy. We philosophize unknowingly, but, as Mario Bunge warns, "The implicit philosophy of the good doctor is that which he practices and not necessarily that which he professes." And it is, at this point, where evaluation is crucial and imperative.[20]

The stable and teaching medical team, in conjunction with the institution, has an undeniable moral and legal responsibility in the evaluation. As well as the State, given that 70% of the medical residencies of our specialty are under its aegis.[21] Evaluation involves *a process of* self-assessment of methodological learning or metacognition strategies. A recent survey by the Argentine Association of Orthopedics and Traumatology, conducted with resident physicians, has identified and alerted the deficit in academic activity as one of the main concerns of young people in training.[21]

The two instruments used in our study, Mini-CEX and DOPS, evaluated "do" or professional excellence. They were chosen *ex profeso*, by virtue of previous experiences that enhance them as of greater educational impact in relation to many others.[8,9,11,15,18] They allowed us to observe the competencies of residents in real situations and various contexts, both clinical and surgical, as well as behavioral factors and horizontal-vertical incorporation of knowledge. At first glance, the observation of the overall results obtained was satisfactory, but the statistical analysis of multivariate regression allowed to identify, precisely, strengths and weaknesses in each year of the residency, as well as to discriminate between different levels of experience. The first corollary for academic leaders of a residency seems obvious: to work to correct these weaknesses. It is also worth noting how misleading descriptive inferences can be in contrast to the power of the scientific method. On the other hand, it is important to note that achieving an optimal skill in one area does not predict success in another, even if they are related. Compared to other publications, we have spent a longer time evaluating surgeries, due to the full observation of procedures.[12,14] However, and similarly to other investigations, there was a high acceptance or satisfaction, both from observers and residents.[11,12,14,18,22,23] Cronbach's α coefficient was 0.90 and 0.92 for DOPS and Mini-CEX, respectively; from a psychometric point of view, factors considered to have the greatest impact on educational measurement.[18] Increasing or sustaining reliability is possible by holding 4 to 10 annual meetings.[9,13,15,22,23] In addition, the increase in observations is also justified by the need to mitigate, statistically, the potential pathology of evaluators: intraobserver variations linked to changes in attention, perspective or mood. We must also recognize interobserver variations: differences in criteria, accuracy, and empathy.[15]

Feedback is the main quality of both instruments, especially *positive feedback,* as it promotes reflection and self-criticism on the part of the resident, helps to identify their weaknesses and strengths. Alves de Lima argues that residents want and need constructive feedback, as they relate them to quality teaching.[24] Feedback must be immediate, since the delay in its implementation leads to evaluation errors. Many researchers, including us, think it is part of our teaching responsibility to influence the professional future of the resident through constant feedback. This is part of socratic teaching, which, despite certain current tensions, we have always professed.[9,22] The time assigned for feedback with Mini-CEX and DOPS was 7 and 8.8 min, respectively, but in future evaluations we will allocate more time, due to the reasons already mentioned. Suggested recommendations for proper feedback are: establishing a respectful learning environment, communicating goals and objectives, comment from direct observation, provide timely feedback on a regular basis, start the session with the self-assessment of the resident, reinforce and correct observed performances, use neutral language to focus on specific knowledge or procedures, confirm understanding and facilitate acceptance, conclude with an action plan, reflect on feedback skills, create opportunities for personal development and make this procedure part of the institutional culture.[18]

This study has several limitations: execution errors related to our first experience with these evaluation tools. On the other hand, given the prevailing culture in our environment, it was unconvincing that the resident did not perceive the review as an exam whose outcome could affect their self-esteem or final qualification. Like other authors, we believe that, given the specific context of each institution, it does not seem advisable to extrapolate, unthinkingly, findings from one residency to another.[23] Finally, we should increase the number of observations in order to sustain or increase the reliability of the evaluation instruments and reduce potential abnormal biases of some observers. The strength of this study was the use of evaluation tools structured in a specific medical residency, inspired by the following thought said by Horace and made known by the excellent German philosopher Immanuel Kant: *Dimidium facti, qui coepit habet, sapere aude. Incipe*.[25]

In conclusion, Mini-CEX and DOPS proved to be relevant and reliable tools for assessing the professionalism of a residency in Pediatric Orthopedics and Traumatology, but this does not imply that they are exclusive. A portfolio that includes various forms of valuation, structured or not, would appear to be the ideal instrument. The quality of the evaluation depends primarily on the relevance to the observed object, the type of prioritized cognitive operations and the subjects involved.

Acknowledgments

Educators who participated in the evaluations: Santiago Beltrán, Fernando Bourdoncle, Sebastián Fedriani Roger, Lucía Molina and Gustavo Viollaz.

Conflict of interests: The authors declare they do not have any conflict of interests.

M. G. Miranda ORCID ID: https://orcid.org/0000-0003-4949-9407

REFERENCES

1. Giacomantone O, Suárez I. Profesionalismo médico, su relación con la educación médica del siglo XXI. Educación Médica Permanente. *Revista de la Sociedad de Educación Médica de La Plata* 2009;I(1):4-18. Available at: http://sedici.unlp.edu.ar/handle/10915/8517
2. Peinado Herreros JM. Competencias médicas. *Educación Médica* 2005;8(Supl. 2):S4-S6. Available at: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-18132005000600002&lng=es
3. Reta-De Rosas AM, López MJ, Vargas AL, Montbrun MC. Evaluación de competencias médicas en un currículo de grado no diseñado por competencias. *Educación Médica* 2008;11(4):207-18. https://doi.org/10.4321/S1575-18132008000400005
4. Tello C. Comité de Residencias. *Rev Asoc Argent Ortop Traumatol* 2015;80(3):143-4. https://doi.org/10.15417/550
5. Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach* 2011;33(3):206-14. https://doi.org/10.3109/0142159X.2011.551559
6. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(Suppl):S63-70. https://doi.org/10.1097/00001888-199009000-00045
7. Miller A, Archer J. Impact of workplace-based assessment on doctors´ education and performance: a systematic review. *BMJ* 2010;341:50-64. https://doi.org/10.1136/bmj.c5064
8. Fornells-Vallés JM. El ABC del Mini-CEX. *Educación Médica* 2009;12(2):83-9. Available at: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-18132009000300004&lng=es
9. Schuwirth LWT, van der Vleuten CPM. Conceptualising surgical education assessment. En: Fry H, Kneebone R (eds.). *Surgical education: Theorising an emerging domain*. Springer Science; 2011:81-9. https://doi.org/10.1007/978-94-007-1682-7

10. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The Mini-CEX (Clinical Evaluation Exercise): a preliminary investigation. *Ann Intern Med* 1995;123:795-9. https://doi.org/10.7326/0003-4819-123-10-199511150-00008
11. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins R. Construct validity of the miniclinical evaluation exercise (miniCEX). *Acad Med* 2003;78(8):826-30. https://doi.org/10.1097/00001888-200308000-00018
12. Amini A, Shirzad F, Mohseni M, Sadeghpour A, Elmi A. Designing Direct Observation of Procedural Skills (DOPS) test for selective skills of orthopedic residents and evaluating its effects from their points of view. *Res Dev Med Educ* 2015;4(2):147-52. https://doi.org/10.15171/rdme.2015.026
13. Beard J, Strachan A, Davies H, Patterson F, Stark P, Ball S, et al. Developing an education and assessment framework for the Foundation Programme. *Med Educ* 2005;39:841-51. https://doi.org/10.1111/j.1365-2929.2005.02236.x
14. Khan AAM, Gorman M, Gwozdziewicz L, Sobani A, Gibson C. Direct Observation of Procedural Skills (DOPS) as an assessment tool for surgical trainees. *J Pak Med Stud* 2013;3(3):137-40. Available at: https://www.jpmsPAGES137-140-PA.pdfonline.com/wp-content/uploads/2019/07/JPMS-VOL3-ISSUE3
15. Durante E. Algunos métodos de evaluación de las competencias: Escalando la pirámide de Miller. *Rev Hosp Ital Buenos Aires* 2006;26(2):55-61. Available at: http://revista.hospitalitaliano.org.ar
16. Nishan FN, Cleland J, McKenzie H, Cassar K. Identifying the factors that determine feedback given to undergraduate medical students following formative Mini-CEX assessments. *Med Educ* 2008;42:89-95. https://doi.org/10.1111/j.1365-2923.2007.02939.x
17. Dawson-Saunders B, Trapp R. *Bioestadística médica*, 2ª ed. México, DF: El Manual Moderno; 1997.
18. Sánchez Mendiola M, González MA. *Evaluación del y para el aprendizaje: instrumentos y estrategias*. Ciudad de México: UNAM; 2020:348. https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es
19. Borges JL. Funes el memorioso. Artificios (1944). En: *Ficciones*, 7ª ed. Buenos Aires: Debolsillo; 2014:125-35.
20. Bunge M. *Filosofía para médicos*. España: Editorial Gedisa; 2012:13-7.
21. Tello C, Vedoya S, Autorino C, Varaona J, Pelaez R, Fazio A, et al. Encuesta dirigida a los residentes sobre aspectos de la Residencia en Ortopedia y Traumatología. *Rev Asoc Argent Ortop Traumatol* 2017;82(2):150-6. https://doi.org/10.15417/694
22. Alves de Lima A, Henquin R, Thierer J. A qualitative study of the impact on learning of the mini-clinical evaluation exercise in postgraduate training. *Med Teach* 2005;27(1):46-52. https://doi.org/10.1080/01421590400013529
23. Fernández Gálvez GM. Evaluación de las competencias clínicas en una residencia de pediatría con el Mini-CEX (Mini-Clinical Evaluation Exercise). *Arch Argent Pediatr* 2011;109(4):314-20. Disponible en: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0325-00752011000400006&lng=es
24. Alves de Lima A. Devolución constructiva una estrategia para mejorar el aprendizaje. *Medicina (Buenos Aires)* 2008;68:88-92. Available at: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0025-76802008000100014
25. Kant I. *Filosofía de la historia*. Buenos Aires: Terramar Ediciones; 2004:33.