

## SEÇÃO: ARTIGOS

# A PROVA PRÁTICA-ORAL ESTRUTURADA É COMPARÁVEL AO EXAME CLÍNICO OBJETIVO ESTRUTURADO NA AVALIAÇÃO DE MICRO-HABILIDADES CLÍNICAS?

Flávia Soares de Matos<sup>1</sup>, Antonio Toledo Jr.<sup>2</sup>

### RESUMO

A simulação tornou-se estratégia de ensino importante em cursos da área da Saúde. O exame clínico objetivo estruturado (OSCE) é considerado padrão ouro na avaliação de habilidades e competências em simulação, mas ele exige muitos recursos, dificultando sua realização frequente. O objetivo deste estudo foi comparar o OSCE com a prova prática-oral estruturada na avaliação de micro-habilidades clínicas. Foi realizada análise quasi-experimental com 21 estudantes de Medicina comparando os dois tipos de prova na avaliação de cinco habilidades obstétricas em ambiente simulado. Os discentes foram divididos em dois grupos. Na primeira fase, cada grupo realizou um tipo de prova. Após 3 semanas, eles realizaram as provas novamente de modo invertido. O tipo de prova (prova prática-oral estruturada ou OSCE) não influenciou a avaliação dos alunos ( $p > 0,05$ ), porém as notas foram mais elevadas ( $p < 0,05$ ) no segundo dia, independentemente do formato de exame, sugerindo efeito teste.

**Palavras-chave:** Educação médica. Avaliação educacional. Competência clínica.

### Como citar este documento – ABNT

MATOS, Flávia Soares de; TOLEDO JR., Antonio. A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?. *Revista Docência do Ensino Superior*, Belo Horizonte, v. 10, e019242, p. 1-21, 2020. DOI: <https://doi.org/10.35699/2237-5864.2020.19242>.

Recebido em: 13/02/2020  
Aprovado em: 13/04/2020  
Publicado em: 15/09/2020

<sup>1</sup> Universidade José do Rosário Vellano em Belo Horizonte (UNIFENAS/BH), Belo Horizonte, MG, Brasil.  
ORCID ID: <https://orcid.org/0000-0001-5724-8128>. E-mail: [flavia.matos@unifenas.br](mailto:flavia.matos@unifenas.br)

<sup>2</sup> Universidade José do Rosário Vellano em Belo Horizonte (UNIFENAS/BH), Belo Horizonte, MG, Brasil.  
ORCID ID: <https://orcid.org/0000-0001-8912-2589>. E-mail: [antonio.toledo@unifenas.br](mailto:antonio.toledo@unifenas.br)

## ¿LA PRUEBA DE PRÁCTICA ORAL ESTRUCTURADA ES COMPARABLE AL EXAMEN CLÍNICO OBJETIVO ESTRUCTURADO EN LA EVALUACIÓN DE MICRO HABILIDADES CLÍNICAS?

### RESUMEN

La simulación se ha convertido en una importante estrategia de enseñanza en los cursos de salud. El examen clínico objetivo estructurado (OSCE) se considera el estándar de oro en la evaluación de habilidades/competencias en la simulación, pero requiere muchos recursos, lo que dificulta su rendimiento frecuente. El objetivo de este estudio fue comparar el OSCE con la prueba de práctica oral estructurada en la evaluación de micro habilidades clínicas en un escenario simulado. Se realizó un estudio cuasiexperimental con 21 estudiantes de medicina que comparó los dos tipos de pruebas en la evaluación de cinco habilidades obstétricas en un escenario simulado. Los estudiantes fueron divididos en dos grupos. En la primera fase, cada grupo realizó un tipo de prueba. Después de 3 semanas, se invirtieron y volvieron a realizar la prueba. El tipo de prueba (OSCE u la prueba de práctica oral estructurada) no influyó en la evaluación de los estudiantes ( $p > 0,05$ ), pero las calificaciones fueron más altas en el segundo día ( $p < 0,05$ ), independientemente del tipo de prueba, sugiriendo un efecto de prueba.

**Palabras clave:** Educación médica. Evaluación educacional. Competencia clínica.

## IS STRUCTURED ORAL PRACTICE TEST COMPARABLE TO OBJECTIVE STRUCTURED CLINICAL EXAMINATION IN THE ASSESSMENT OF CLINICAL MICRO-SKILLS?

### ABSTRACT

Simulation has become an important teaching strategy in health courses. The objective structured clinical examination (OSCE) is considered a gold standard in the assessment of skills and competences in simulation, but it requires many resources, making it difficult to perform it frequently. The objective of this study was to compare the OSCE with the structured oral-practical test in the assessment of clinical micro-skills. A quasi-experimental analysis was carried out with 21 medical students comparing the two types of test in the evaluation of five obstetric skills in a simulated environment. The students were divided into two groups. In the first phase, each group performed a type of test. After three weeks, the groups were switched and students took the tests again. The test type (OSCE or structured oral-practice test) did not influence the students' assessment ( $p > 0.05$ ). However, the scores were higher on the second day ( $p < 0.05$ ) regardless of the exam format, suggesting a "test effect".

**Keywords:** Medical education. Educational evaluation. Clinical competence.

## INTRODUÇÃO

O ensino de competências clínicas por meio da simulação é uma inovação nos currículos das escolas médicas. Ele permite o treinamento dos estudantes de forma ética, segura e sem riscos para os pacientes. O termo competências clínicas envolve habilidades como coleta da anamnese, realização do exame físico e de procedimentos, raciocínio clínico, comunicação, trabalho em equipe e profissionalismo. O laboratório de simulação oferece ambiente protegido e seguro para alunos e professores, onde se pode treinar habilidades e competências clínicas de maneira repetida, sem constrangimento, reduzindo a ansiedade, antes de praticá-las em cenário real, assim o discente poderá abordar o paciente com mais confiança futuramente (AL-ELQ, 2007; RAZAVI *et al.*, 2010). No entanto, essa nova estratégia de ensino necessita de novas formas de avaliação.

Durante muito tempo, o sistema de avaliação das escolas médicas foi dominado por provas voltadas apenas para avaliação do conhecimento cognitivo. Até a década de 1970, a avaliação de competências clínicas era feita por métodos tradicionais, como provas de múltipla escolha, pequenos ensaios e questões abertas, adequados para a avaliação do conhecimento exclusivamente cognitivo. A competência clínica era avaliada por meio da discussão de casos à beira do leito dos pacientes, limitada aos aspectos teóricos das doenças, e sem nenhuma demonstração prática por parte do aluno. Além disso, só eram abordadas as doenças de pacientes internados no momento da avaliação. Apesar do exame físico ser considerado de suma importância na avaliação da competência clínica, muitas vezes ele era feito por meio de fotos, sem que houvesse a demonstração completa das habilidades técnicas (HARDEN; GLEESON, 1979; HARDEN *et al.*, 1975). As avaliações acadêmicas falhavam em documentar o comportamento dos estudantes frente a frente com o paciente. Habilidades profissionais práticas, como colher a história clínica, realizar o exame físico, solicitar exames complementares, analisar e interpretar essas informações, formular hipóteses diagnósticas e elaborar plano terapêutico, não eram avaliadas objetivamente (HARDEN; GLEESON, 1979). A avaliação somativa de competências clínicas deveria ser feita com provas práticas, envolvendo tarefas clínicas, capazes de avaliar o desempenho profissional (MILLER, 1990; VAN DER VLEUTEN *et al.*, 2010; WASS *et al.*, 2001). A ideia de que nenhum método de avaliação isolado seria capaz de mensurar algo tão complexo como as competências médicas levou à proposta de associação de estratégias de avaliação (ELDER, 2018; GORMLEY, 2011; HARDEN; GLEESON, 1979; MILLER, 1990; THIESSEN; FISCHER; HUVENDIEK, 2019; TIJANI *et al.*, 2017).

A prova oral, uma das formas tradicionais de avaliação mais utilizadas, não é estruturada quanto ao número, tipo, pontuação e tempo de execução das questões. O aluno fica diante do avaliador, dessa maneira sua compreensão sobre a prática clínica pode ser explorada

(ANASTAKIS; COHEN; REZNICK, 1991; SMEE, 2003). O teste pode tomar a forma de entrevista ou de discussão e reflete o modo de comunicação dominante na prática profissional. Ele pode testar os limites do conhecimento e da compreensão do estudante e avaliar algumas habilidades, como raciocínio clínico, competência para resolver problemas, e qualidades, como confiança e segurança. Por outro lado, a prova apresenta dificuldades, como limitação no número de habilidades testadas, decisões tomadas de forma independente pelo avaliador a partir de critérios desconhecidos, impossibilidade de recurso por parte do discente, dificuldade de diferenciar o quê o aluno fala, de como ele fala e má performance devido ao estresse. Esse tipo de prova é muito utilizado nas ciências básicas, como Anatomia e Patologia. Nela, um ou mais avaliadores fazem de cinco a nove perguntas para o avaliado, com duração de 10 a 20 minutos. Seu uso é criticado em função das baixas confiabilidade e validade, devidas, em parte, ao papel ativo do avaliador, o qual pode aumentar, ou até mesmo reduzir, o tempo, o número, ou a profundidade das questões, influenciado por aparência, preconceito ou pelo próprio desempenho inicial do estudante, conhecido como efeito halo (DAVIS; KARUNATHILAKE, 2005; JAYAWICKRAMARAJAH, 1985; WASS *et al.*, 2003).

Havia então a necessidade de uma prova com maior validade e confiabilidade, para melhorar a avaliação. Em 1975, Harden propôs um novo método, chamado de exame clínico estruturado. Essa prova continha 16 estações avaliativas, sendo oito procedimentais, intercaladas com mais oito contendo questões curtas, ou de múltipla escolha, sobre os procedimentos realizados. Cada estação tinha a duração de 5 minutos. As competências clínicas do estudante seriam avaliadas isoladamente em cada estação, pela demonstração de habilidades em pacientes simulados (HARDEN *et al.*, 1975). Ela foi realizada pela primeira vez em 1977, na *Dundee Medical School*, como piloto, com 20 estações de 4 ou 5 minutos de duração e avaliava competências cirúrgicas (DAVIS, 2003). A partir de 1978, passou a ser realizada rotineiramente pela clínica cirúrgica, ao final de um período letivo e, posteriormente, estendeu-se a todas as áreas (DAVIS, 2003). Esse tipo de avaliação foi denominada *Objective Structured Clinical Examination* ou exame clínico objetivo estruturado (OSCE).

Nas últimas 4 décadas, o OSCE cresceu exponencialmente em todo o mundo e, atualmente, é considerado padrão ouro para avaliação de competências clínicas pelas principais associações de ensino médico (HARDEN, 2016; KHAN, K. *et al.*, 2013b). É utilizado na maioria das escolas médicas dos Estados Unidos, do Reino Unido e do Canadá como ferramenta para avaliar os alunos da graduação ao término de um período letivo (HARDEN, 2015, 2016; KHAN, K. *et al.*, 2013b). É uma prova de alto custo, estruturada em rodízio de estações clínicas, com pacientes padronizados, sendo o avaliador de cada estação previamente treinado com *checklist* específico, de modo a padronizar a avaliação e aumentar sua

confiabilidade. Trata-se de uma prova flexível, na qual os estudantes são avaliados em 6 a 25 estações com duração de 4 a 20 minutos cada. As competências clínicas são divididas e distribuídas por estação (HARDEN, 1988; KHAN, K. *et al.*, 2013a; SMEE, 2003). O exame envolve longo planejamento prévio, grande número de professores e de materiais e área física adequada (MILLER, 1990; SMEE, 2003).

Apesar das provas do tipo prática-oral convencionais não estruturadas terem sua validade e confiabilidade menores quando comparadas ao OSCE, ainda são muito utilizadas em todo o mundo, inclusive no Brasil (ANASTAKIS; COHEN; REZNICK, 1991; EPSTEIN, 2007). Alguns estudos propõem que a prova prática-oral tem sua confiabilidade aumentada quando sua finalidade é avaliar competências práticas, quando tem seu conteúdo padronizado com utilização de *checklist* e quando se utiliza avaliadores treinados e focados nas habilidades profissionais (MEMON; JOUGHIN; MEMON, 2010; SMEE, 2003; WASS *et al.*, 2003). O alto custo e complexidade do OSCE, aliado à tradição, à simplicidade e ao baixo custo da prova oral, levou algumas instituições a modificarem a estrutura do sistema de avaliação, implementando uma prova prática-oral estruturada, com número, tipo, pontuação e tempo de execução das questões predeterminados (ANASTAKIS; COHEN; REZNICK, 1991; HASHIM *et al.*, 2015).

Estudos comparando o OSCE com provas tradicionais, como múltipla escolha ou apresentação de casos longos e curtos, foram realizados com estudantes de Medicina, Enfermagem, Odontologia, Psicologia, Fonoaudiologia, Fisioterapia, Farmácia e Veterinária. Nos questionários de satisfação aplicados, eles preferiram o OSCE (AL OMARI; SHAWAGFA, 2010; AMIRI; NIKBAKHT, 2012; MAJUMDER *et al.*, 2019; MONDAL *et al.*, 2012). Ainda assim, no estudo realizado por Al Omari e Shawagfa (2010), os discentes acharam mais fácil lidar com pacientes simulados, mas consideraram a prova mais estressante e mais difícil em comparação com a prova tradicional. Em pesquisa realizada por Pierre *et al.* (2004), igualmente, os alunos consideraram o OSCE mais estressante e acharam que era preciso mais tempo para a resolução da tarefa. O tempo também foi um fator de insatisfação no trabalho de Amiri e Nikbakht (2012).

Um estudo do Departamento de Pediatria do *Institute of Post Graduate Medical Education and Research*, Calcutá, Índia, comparou o OSCE com a prova tradicional, dos tipos casos longos, curtos e prática-oral, em relação a estudantes do nono semestre. As avaliações foram realizadas em dois dias consecutivos, sendo um dia para a prova tradicional e outro para o OSCE. Após a segunda prova, foi distribuído um questionário para avaliar a percepção dos discentes sobre os dois tipos de teste. Seu desempenho foi melhor no OSCE em comparação com a prova tradicional. A maior parte deles, 73,8%, preferiu o OSCE (MONDAL *et al.*, 2012). Um estudo do Departamento de Pediatria do *Medical College*, Thrissur, Índia,

comparou o micro-OSCE, realizado com estudantes do quinto semestre, com o OSCE completo, realizado com alunos do oitavo semestre. O micro-OSCE continha duas estações, sendo uma delas de coleta de história clínica e, a outra, de exame físico em pediatria, com *checklist* definido. O desempenho dos discentes foi melhor no micro-OSCE, provavelmente devido à menor quantidade de conteúdos avaliados (MATHEWS; MENON; MANI, 2004). Um estudo da Universidade de Lagos, Nigéria, comparou avaliações dos tipos casos longos, ensaios, questões de múltipla escolha e OSCE com estudantes da clínica cirúrgica ao longo de 4 anos para avaliar qual era o melhor tipo de prova. O desempenho foi melhor no OSCE, porém a associação deste com casos longos apresentou maior poder de avaliação (TIJANI *et al.*, 2017).

Existem poucas pesquisas comparando a prova prática-oral estruturada com o OSCE. Lunenfeld *et al.* (1991) compararam um OSCE com sete estações de 5 minutos pareadas com estações de prova prática-oral estruturada em Israel. O objetivo era avaliar se o OSCE detectava mais deficiências do aluno. A diferença básica entre as provas era o formato da estação e a interação com o avaliador. De acordo com os resultados, o OSCE foi mais difícil e foi capaz de detectar mais deficiências entre os estudantes. Em estudos comparando a prova prática-oral não estruturada com a prova prática-oral estruturada, o desempenho dos discentes foi melhor na prova estruturada (HASHIM *et al.*, 2015; KHAN, H.; MIRZA, 2017).

Apesar de o OSCE ser reconhecido como padrão ouro para avaliação de competências e habilidades clínicas, seu custo e complexidade de planejamento e operacional dificultam sua aplicação repetidas vezes ao longo do período letivo. Diante disso, torna-se necessária a validação de métodos de avaliação menos complexos, que possam ser realizados mais vezes ao longo do semestre e sejam complementares ao OSCE. O objetivo desta análise foi comparar a prova prática-oral estruturada com uma estação tipo OSCE na avaliação de micro-habilidades clínicas de estudantes de Medicina.

## MATERIAIS E MÉTODOS

### Desenho e população

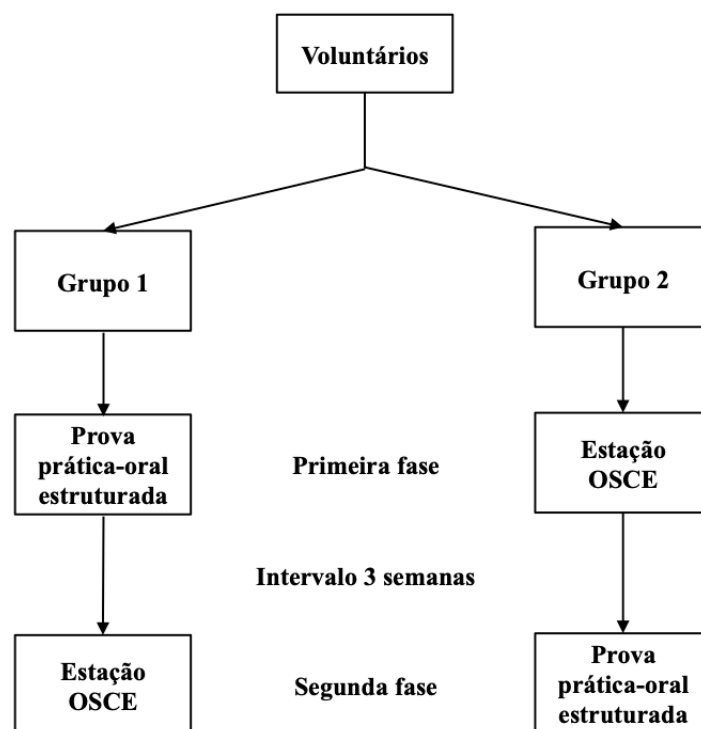
Este é um estudo prospectivo e quasi-experimental para comparação de dois métodos de avaliação de micro-habilidades clínicas entre estudantes de Medicina. Todos os 66 alunos matriculados no quarto período do curso de Medicina da Universidade José do Rosário Vellano, campus de Belo Horizonte, no primeiro semestre de 2017 foram convidados a participar da pesquisa. Foi utilizada uma amostra não probabilística de conveniência.

Os critérios de inclusão foram: estar regularmente matriculado no curso, ter sido aprovado no bloco Gestação até o segundo semestre de 2016 e assinar o Termo de Consentimento

Livre e Esclarecido (TCLE). Os critérios de exclusão foram: situação acadêmica irregular, ou não ter cursado o bloco Gestação, ou já ter cursado o bloco Síndromes Ginecológicas e Obstétricas, ou não assinar o TCLE.

### Procedimentos

Os voluntários foram recebidos numa sala, onde foi explicado o projeto, o TCLE e foram esclarecidas dúvidas. Após a assinatura do TCLE, eles foram distribuídos em dois grupos e submetidos a dois tipos de prova. Na primeira fase da pesquisa, o Grupo 1 foi submetido à prova prática-oral estruturada (POE) e o Grupo 2 realizou a prova em estação tipo OSCE. Para evitar a mudança do tipo de prova a cada voluntário, o Grupo 1 realizou a prova primeiramente, seguido do Grupo 2. Enquanto esperavam sua vez de fazer a prova, os alunos dos dois grupos ficaram confinados na sala para evitar contaminação. Três semanas depois, foi realizada a segunda fase do estudo. Os estudantes foram confinados novamente e os grupos realizaram as provas de modo invertido (FIGURA 1).



**Figura 1** – Desenho do estudo  
Fonte: elaborada pelos autores, 2020.

Os testes avaliaram cinco habilidades obstétricas: três primeiras manobras de Leopold (localização do fundo uterino, identificação da situação e posição fetal e determinação da apresentação fetal), medida de útero-fita e ausculta de batimentos cardíacos do feto em

A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?

Flávia Soares de Matos, Antonio Toledo Jr.

manequim. Na POE, os alunos realizaram as manobras em sequência, uma a uma, após comando verbal do aplicador para cada uma delas (QUADRO 1). As cinco questões foram realizadas no prazo máximo de 4 minutos e os discentes puderam interagir com o avaliador. Na prova tipo OSCE, os estudantes analisaram, por 1 minuto, a descrição de uma situação clínica ao entrar na sala (QUADRO 1) e tiveram 4 minutos para realizar a prova. Não foi permitido interação com o avaliador.

---

#### **Prova tipo prática-oral estruturada**

---

Ordem verbal para o aluno:

Demonstre a medida de útero fita e a ausculta de batimentos cardíacos fetais no manequim.

---

#### **Estação tipo OSCE**

---

Instrução para o aluno:

Você se encontra em uma Unidade Básica de Saúde e está atendendo uma gestante de risco habitual.

Dados da anamnese: paciente de 28 anos, G1P0A0, DUM 15/11/2016, sem história de dor, sangramento ou perda de líquido. Movimentação fetal presente.

Exames laboratoriais sem alterações.

Você deverá realizar o exame obstétrico dessa paciente.

---

**Quadro 1** – Questão utilizada na prova prática-oral estruturada e na estação tipo OSCE

Fonte: elaborado pelos autores, 2020.

Todas as provas foram aplicadas por um único avaliador. Foi utilizado o mesmo *checklist* (QUADRO 2) para atribuição da nota nos dois tipos de prova. O valor total da prova foi de 26 pontos. A pontuação de cada questão foi baseada no número de micro-habilidades necessárias para a realização da tarefa. Foram distribuídos dois pontos para a realização correta de cada micro-habilidade e um ponto para sua realização parcialmente correta.



A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?

Flávia Soares de Matos, Antonio Toledo Jr.

| Questão  | Como realizou a manobra? |                |           |
|--|--------------------------|----------------|-----------|
|  | Não (0,0)                | Razoável (1,0) | Sim (2,0) |
| <b>1. Primeira manobra de Leopold (4,0 pontos)</b>                                     |                          |                |           |
| • Realizou a manobra corretamente?   |                          |                |           |
| • Localizou corretamente o fundo uterino?  |                          |                |           |
| <b>2. Medida de altura uterina (6,0 pontos)</b>  |                          |                |           |
| • Posicionou a fita métrica corretamente?  |                          |                |           |
| • Posicionou as mãos corretamente?   |                          |                |           |
| • Leu corretamente a medida do fundo uterino?  |                          |                |           |
| <b>3. Segunda manobra de Leopold (6,0 pontos)</b>                                      |                          |                |           |
| • Realizou a manobra corretamente?   |                          |                |           |
| • Falou corretamente a situação fetal?   |                          |                |           |
| • Falou corretamente a posição fetal?  |                          |                |           |
| <b>4. Terceira manobra de Leopold (4,0 pontos)</b>                                     |                          |                |           |
| • Realizou a manobra corretamente?   |                          |                |           |
| • Falou corretamente a apresentação fetal?   |                          |                |           |
| <b>5. Ausculta dos batimentos cardíacos fetais (6,0 pontos)</b>                        |                          |                |           |
| • Posicionou o estetoscópio de Pinard no dorso do feto, mais próximo ao polo cefálico? |                          |                |           |
| • Retirou a mão do estetoscópio?   |                          |                |           |
| • Comparou o batimento do feto com o pulso da gestante?                                |                          |                |           |

**Quadro 2** – Checklist utilizado para avaliação das duas provas

Fonte: elaborado pelos autores, 2020.

### Plano de análise estatística

O desfecho primário foi a nota em cada uma das provas. A variável principal foi o tipo de prova e as variáveis independentes foram: sexo, idade, aproveitamento global no curso (média aritmética das notas dos módulos cursados), percentual de aproveitamento no Treinamento de Habilidades (TH) do bloco Gestação e se o aluno foi monitor de TH. Para fins de análise, utilizou-se a nota de cada uma das cinco questões isoladamente e a nota total da prova.

A análise descritiva foi feita por meio da distribuição de frequência para as variáveis categóricas e média e desvio padrão para as variáveis contínuas. A análise comparativa teve o tipo de prova como variável principal e a nota da prova como desfecho. Foi utilizado o  $\chi^2$ , ou teste exato de Fisher, para comparação das variáveis categóricas e o teste t de Student para as variáveis contínuas. Comparou-se a nota em cada questão e a nota total nos seguintes cruzamentos: tipo de prova em cada uma das fases do estudo; tipo de prova independentemente da fase; OSCE versus POE nas duas fases e OSCE na primeira fase versus POE na segunda fase e vice-versa. A análise multivariada foi realizada por meio de regressão

linear. Foi considerado nível de significância de 0,05. Utilizou-se o software IBM® SPSS® Statistics v. 19 (IBM, USA) na análise estatística.

### Aspectos éticos

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade José do Rosário Vellano (Parecer nº 2.048.558) e está em concordância com a versão atual da Declaração de Helsinque e a Resolução 466/2012 do CONEP e suas atualizações. Todos os voluntários receberam esclarecimento sobre os objetivos da pesquisa e assinaram o TCLE antes de sua inclusão.

### RESULTADOS

26 alunos do quarto período do curso de Medicina da Universidade José do Rosário Vellano, campus de Belo Horizonte, foram recrutados. Um aluno foi excluído por estar em situação acadêmica irregular e quatro deles foram excluídos por terem participado apenas da primeira fase do estudo. O Grupo 1 foi formado por 13 e o Grupo 2 por 8 sujeitos, totalizando 21 estudantes, correspondendo a 31,8% da população total. A primeira fase foi realizada no dia 2 de junho de 2017 e, a segunda, em 23 de junho de 2017.

A maioria dos alunos era do sexo feminino (61,9%) e quatro eram monitores do TH (19,0%). A idade média foi de  $21,2 \pm 2,4$  anos. O aproveitamento global médio foi de  $82,5 \pm 5,5$  pontos e o percentual de aproveitamento médio do TH no bloco Gestação foi de  $87,2 \pm 7,9\%$ . A Tabela 1 detalha os dados demográficos e desempenho acadêmico dos voluntários. Não houve diferença entre os grupos, exceto no aproveitamento global, que foi maior no Grupo 2 ( $85,7 \pm 5,8$  versus  $80,6 \pm 4,4$ , respectivamente Grupo 2 e Grupo 1).

| Variável                            | Grupo 1<br>n=13 |     | Grupo 2<br>n=8 |     | Total<br>n=21 |     | P     |
|-------------------------------------|-----------------|-----|----------------|-----|---------------|-----|-------|
|                                     | Média           | DP  | Média          | DP  | Média         | DP  |       |
| Idade (anos)                        | 21,0            | 0,9 | 21,5           | 3,9 | 21,2          | 2,4 | 0,655 |
| Mulheres (%)                        | 76,9            | --- | 37,5           | --- | 61,9          | --- | 0,164 |
| Aproveitamento global no curso      | 80,6            | 4,4 | 85,7           | 5,8 | 82,5          | 5,5 | 0,031 |
| Percentual de aproveitamento TH (%) | 88,4            | 8,5 | 85,3           | 6,9 | 87,2          | 7,9 | 0,386 |

TH - treinamento de habilidades; DP - desvio padrão.

**Tabela 1** – Dados demográficos e desempenho acadêmico prévio dos voluntários

Fonte: elaborada pelos autores, 2020.

A Tabela 2 compara as notas médias por questão e a nota final de acordo com o tipo e a fase de realização da prova. A nota total média para OSCE e para POE foi, respectivamente,  $19,6 \pm 3,7$  e  $20,0 \pm 3,8$ , na primeira fase, e  $23,5 \pm 2,8$  e  $23,4 \pm 2,6$ , na segunda fase, para um total

de 26 pontos. Não houve diferença entre as notas das questões e a nota total quando se comparou os diferentes tipos de prova em cada uma das fases. Também não se observou diferença quando se comparou as notas por tipo de prova, independentemente do grupo ou da fase do estudo. Nesse caso a nota total média do OSCE foi 22,1±3,6 em comparação com 21,3±3,7 na POE ( $p=0,453$ ).

| Fase 1  | Grupo 1 (POE)<br>n=13 |     | Grupo 2 (OSCE)<br>n=8 |     | p     |
|---|-----------------------|-----|-----------------------|-----|-------|
|   | Média                 | DP  | Média                 | DP  |       |
| Questão 1 – 1ª manobra de Leopold                   | 2,8                   | 1,6 | 2,8                   | 1,8 | 0,980 |
| Questão 2 – medida de útero-fita                    | 3,9                   | 1,6 | 4,0                   | 1,8 | 0,841 |
| Questão 3 – 2ª manobra de Leopold                   | 4,9                   | 1,1 | 4,8                   | 1,0 | 0,848 |
| Questão 4 – 3ª manobra de Leopold                   | 3,9                   | 0,3 | 3,6                   | 0,7 | 0,203 |
| Questão 5 – ausculta de batimentos cardíacos fetais | 4,6                   | 1,3 | 4,5                   | 1,4 | 0,848 |
| Nota total  | 20,0                  | 3,8 | 19,6                  | 3,7 | 0,826 |
| Fase 2  | Grupo 1 (OSCE)        |     | Grupo 2 (POE)         |     | p     |
|   | Média                 | DP  | Média                 | DP  |       |
| Questão 1 – 1ª manobra de Leopold                   | 3,5                   | 1,1 | 3,8                   | 0,5 | 0,622 |
| Questão 2 – medida de útero-fita                    | 5,4                   | 0,7 | 5,1                   | 1,7 | 0,627 |
| Questão 3 – 2ª manobra de Leopold                   | 5,3                   | 1,0 | 5,3                   | 1,2 | 0,902 |
| Questão 4 – 3ª manobra de Leopold                   | 3,9                   | 0,3 | 4,0                   | 0,0 | 0,447 |
| Questão 5 – ausculta de batimentos cardíacos fetais | 5,4                   | 1,0 | 5,3                   | 1,0 | 0,765 |
| Nota total  | 23,5                  | 2,8 | 23,4                  | 2,6 | 0,894 |

POE – prova prática-oral estruturada; OSCE – exame clínico objetivo estruturado; DP – desvio padrão.

**Tabela 2** – Comparação entre as notas médias obtidas pelos grupos de acordo com o tipo de prova e a fase do estudo

Fonte: elaborada pelos autores, 2020.

A Tabela 3 compara as notas em cada um dos grupos de acordo com o tipo de prova. No Grupo 1 as notas do OSCE foram maiores comparadas com as da POE, na questão 2 e na nota total. No Grupo 2 observa-se o fenômeno contrário, as notas da POE foram maiores, quando comparadas com as do OSCE, na questão 2 e na nota total. Em ambos os grupos, os tipos de prova com maior nota correspondem à prova realizada na segunda fase.

A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?

Flávia Soares de Matos, Antonio Toledo Jr.

| Grupo 1   | POE (Fase 1)<br>n=13 |     | OSCE (Fase 2)<br>n=13 |     | p*           |
|---|----------------------|-----|-----------------------|-----|--------------|
|   | Média                | DP  | Média                 | DP  |              |
| Questão 1 – 1ª manobra de Leopold                   | 2,8                  | 1,6 | 3,5                   | 1,1 | 0,054        |
| Questão 2 – medida de útero-fita                    | 3,9                  | 1,6 | 5,4                   | 0,7 | <b>0,002</b> |
| Questão 3 – 2ª manobra de Leopold                   | 4,9                  | 1,1 | 5,3                   | 1,0 | 0,190        |
| Questão 4 – 3ª manobra de Leopold                   | 3,9                  | 0,3 | 3,9                   | 0,3 | 1,000        |
| Questão 5 – ausculta de batimentos cardíacos fetais | 4,6                  | 1,3 | 5,4                   | 1,0 | 0,054        |
| Nota total  | 20,0                 | 3,8 | 23,5                  | 2,8 | <b>0,001</b> |

| Grupo 2   | POE (Fase 2)<br>n=8 |     | OSCE (Fase 1)<br>n=8 |     | p*           |
|---|---------------------|-----|----------------------|-----|--------------|
|   | Média               | DP  | Média                | DP  |              |
| Questão 1 – 1ª manobra de Leopold                   | 3,8                 | 0,5 | 2,8                  | 1,8 | 0,121        |
| Questão 2 – medida de útero-fita                    | 5,1                 | 1,7 | 4,0                  | 1,8 | <b>0,007</b> |
| Questão 3 – 2ª manobra de Leopold                   | 5,3                 | 1,2 | 4,8                  | 1,0 | 0,430        |
| Questão 4 – 3ª manobra de Leopold                   | 4,0                 | 0,0 | 3,6                  | 0,7 | 0,197        |
| Questão 5 – ausculta de batimentos cardíacos fetais | 5,3                 | 1,1 | 4,5                  | 1,4 | 0,197        |
| Nota total  | 23,4                | 3,7 | 19,6                 | 2,6 | <b>0,005</b> |

POE – prova prática-oral; OSCE – exame clínico objetivo estruturado; DP – desvio padrão; \* teste de t de Student pareado.

**Tabela 3** – Comparação entre as notas de POE e OSCE em cada grupo

Fonte: elaborada pelos autores, 2020.

Para testar a hipótese da influência da fase do estudo no desempenho dos alunos, compararam-se as notas da primeira fase com as notas da segunda fase, independentemente do tipo de prova (TABELA 4). As notas da segunda fase foram maiores em comparação com as notas da primeira fase. Houve diferença nas questões 1, 2 e 5 e na nota total. Destaca-se que a nota da questão quatro foi elevada nas duas fases da pesquisa.

A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?

Flávia Soares de Matos, Antonio Toledo Jr.

| Variável  | Primeira fase<br>n=21 |     | Segunda fase<br>n=21 |     | p*               |
|---|-----------------------|-----|----------------------|-----|------------------|
|   | Média                 | DP  | Média                | DP  |                  |
| Questão 1 – 1ª manobra de Leopold                   | 2,8                   | 1,7 | 3,6                  | 0,9 | <b>0,010</b>     |
| Questão 2 – medida de útero-fita                    | 3,9                   | 1,6 | 5,3                  | 1,2 | <b>&lt;0,001</b> |
| Questão 3 – 2ª manobra de Leopold                   | 4,8                   | 1,1 | 5,3                  | 1,0 | 0,125            |
| Questão 4 – 3ª manobra de Leopold                   | 3,8                   | 0,5 | 4,0                  | 0,2 | 0,267            |
| Questão 5 – ausculta de batimentos cardíacos fetais | 4,6                   | 1,3 | 5,3                  | 1,0 | <b>0,017</b>     |
| Nota total  | 19,9                  | 3,7 | 23,5                 | 2,6 | <b>&lt;0,001</b> |

DP - desvio padrão; \*teste de t de Student pareado.

**Tabela 4** – Comparação entre as notas da primeira e da segunda fase independentemente do tipo de prova

Fonte: elaborada pelos autores, 2020.

Realizou-se então análise multivariada por meio de regressão linear para avaliar qual dos dois fatores, o tipo de prova ou a fase do estudo, era determinante na nota da avaliação. A Tabela 5 demonstra associação significativa entre a fase do estudo e a nota total e das questões 1, 2 e 5, comprovando as observações da Tabela 4.

| <b>Questão 1 – 1ª manobra de Leopold</b>                   | <b>Beta</b> | <b>p</b>     |
|--|-------------|--------------|
| Tipo de prova  | 0,042       | 0,792        |
| Fase do estudo   | 0,320       | <b>0,048</b> |
| <b>Questão 2 – medida de útero-fita</b>                    | <b>Beta</b> | <b>p</b>     |
| Tipo de prova  | -0,067      | 0,651        |
| Fase do estudo   | 0,431       | <b>0,006</b> |
| <b>Questão 3 – 2ª manobra de Leopold</b>                   | <b>Beta</b> | <b>p</b>     |
| Tipo de prova  | 0,009       | 0,955        |
| Fase do estudo   | 0,230       | 0,160        |
| <b>Questão 4 – 3ª manobra de Leopold</b>                   | <b>Beta</b> | <b>p</b>     |
| Tipo de prova  | 0,240       | 0,135        |
| Fase do estudo   | 0,240       | 0,135        |
| <b>Questão 5 – ausculta de batimentos cardíacos fetais</b> | <b>Beta</b> | <b>p</b>     |
| Tipo de prova  | -0,004      | 0,979        |
| Fase do estudo   | 0,323       | <b>0,045</b> |
| <b>Nota total</b>  | <b>Beta</b> | <b>p</b>     |
| Tipo de prova  | 0,015       | 0,918        |
| Fase do estudo   | 0,508       | <b>0,001</b> |

**Tabela 5** – Regressão linear entre tipo de prova, fase do estudo e nota das questões, nota total da avaliação

Fonte: elaborada pelos autores, 2020.

## DISCUSSÃO

O objetivo desta análise foi comparar o desempenho de alunos de Medicina em avaliação de micro-habilidades clínicas por prova prática-oral estruturada e por estação tipo OSCE. Não se observou nenhuma diferença entre as notas dos dois tipos de prova na comparação entre os grupos nas duas fases do estudo. Vários fatores, como o pequeno tamanho da amostra, a utilização de amostra de conveniência, a inclusão de estudantes de apenas um período e a ausência de acompanhamento entre as duas fases do estudo, que levou à assimetria entre os grupos, podem ter contribuído para a diminuição do poder da pesquisa em identificar diferenças entre os dois tipos de teste.

Um ponto que poderia ser uma vantagem, e ao mesmo tempo uma desvantagem, foi a semelhança entre as provas, cuja diferença principal foi a forma de aplicação. A prova prática-oral estruturada foi elaborada de acordo com os parâmetros recomendados na literatura (tempo de realização predeterminado, questões predefinidas, *checklist* específico e avaliador previamente treinado) com o objetivo de aumentar a sua validade e confiabilidade (KHAN, H.; MIRZA, 2017; MEMON; JOUGHIN; MEMON, 2010; SMEE, 2003). A estação de OSCE apresentava caso clínico contextualizado e instruções escritas para os alunos. No primeiro tipo de prova houve interação entre o avaliador e os avaliados, por meio da indicação das manobras a serem realizadas, e na segunda prova não houve interação, como recomendado na literatura. Nos dois tipos de avaliação utilizou-se o mesmo *checklist*. Essa uniformidade reduz eventuais vieses associados à execução do projeto, garantindo que as eventuais discrepâncias estejam relacionadas à diferença na aplicação das provas. Ela não foi observada em outros estudos. Por outro lado, essa invariabilidade, associada ao pequeno número de questões, pode ter reduzido a capacidade da pesquisa em identificar eventuais diferenças entre as provas, como observado no estudo de Lunenfeld *et al.* (1991), no qual o desempenho dos discentes foi pior no OSCE.

Apesar dessas limitações, esta análise apresenta também pontos positivos que favorecem a generalização de seus resultados. Trata-se de estudo prospectivo, controlado e com o mesmo critério de avaliação entre os grupos (avaliador e *checklist*). Todos os voluntários cursavam o mesmo período e tinham feito a disciplina sobre o conteúdo da prova na mesma época, reduzindo possíveis vieses de memória e garantindo maior homogeneidade entre os grupos. Além disso, os dois grupos se mostraram comparáveis em relação aos aspectos sociodemográficos e ao desempenho acadêmico. Para avaliar se o maior aproveitamento global no curso apresentado pelo Grupo 2 interferiu na avaliação, comparou-se o desempenho entre os dois grupos considerando-se as duas provas, independentemente da fase do estudo, e não se observou diferença entre eles. Ainda que a amostra tenha sido pequena, utilizou-se teste estatístico pareado na comparação intragrupo, o que aumentou o

poder estatístico da análise. Destaca-se também que os valores de p foram muito pequenos, principalmente em relação à medida de útero fita e à nota final, indicando que as diferenças observadas devem realmente existir.

Embora haja na literatura estudos comparando o OSCE com outros tipos de prova, como questões de múltipla escolha, ensaios e apresentação de casos longos e curtos (AL OMARI; SHAWAGFA, 2010; AMIRI; NIKBAKHT, 2012; MAJUMDER *et al.*, 2019; MONDAL *et al.*, 2012; TIJANI *et al.*, 2017), entende-se que esses outros tipos de prova são mais adequados para a avaliação de conteúdo cognitivo e de habilidades cognitivas e não para a avaliação de habilidades e competências clínicas, que envolvem domínios que não podem ser avaliados exclusivamente por questões escritas.

Um ponto importante a ser considerado na comparação desses dois tipos de prova é o nível de ansiedade e estresse do aluno, que pode interferir diretamente em seu desempenho (YOUNG *et al.*, 2014). Apesar dos relatos de redução do estresse e da ansiedade associados à utilização de manequins e pacientes simulados no ensino de habilidades (AL-ELQ, 2007; RAZAVI *et al.*, 2010) e no OSCE (AL OMARI; SHAWAGFA, 2010), outras pesquisas indicam que os alunos acharam esse tipo de prova mais estressante (BRAND; SCHOONHEIM-KLEIN, 2009; LEBLANC; BANDIERA, 2007). Em dois estudos os discentes associaram o estresse à limitação de tempo para realização da tarefa ou questão (PIERRE *et al.*, 2004; AMIRI; NIKBAKHT, 2012). Isso pode ser agravado em adaptações do OSCE, como o mini-OSCE, que podem levar à redução da duração da questão. Um estudo realizado no Brasil por Troncon (2004), com 258 estudantes de Medicina ao longo de 3 anos, mostrou que a dificuldade de lidar com o estresse e de gerenciamento de tempo foram queixas constantes ao longo do período estudado, não obstante a melhora na avaliação da organização da prova em si. Talvez o nível de estresse possa ser menor na prova prática-oral estruturada, devido à possibilidade de interação com o professor e à menor pressão do limite de tempo, por não se tratar de uma sequência de estações. Esses fatores, teoricamente, favoreceriam a performance do aluno. No entanto, não foram identificados na literatura estudos avaliando essas hipóteses.

O principal achado deste trabalho foi o aumento do desempenho dos alunos na segunda fase do estudo, independentemente do tipo de prova. Observou-se aumento na nota média total e na nota média das questões 1, 2 e 5 na segunda fase. Nas questões 3 e 4, não houve diferença, possivelmente pelo desempenho já elevado na primeira fase. Na questão 3 (segunda manobra de Leopold), o aumento observado entre as duas fases foi de apenas 0,5 ponto, não alcançando significância estatística. Já na questão 4 (terceira manobra de Leopold), os avaliados apresentaram desempenho inicial alto, média de 3,8 pontos num total possível de 4,0, considerando os dois tipos de prova, indicando efeito teto, impeditivo de melhora significativa da nota na segunda fase.

Como foram avaliadas as mesmas habilidades nas duas fases da pesquisa, esses achados sugerem a presença de efeito teste, apesar do intervalo de 3 semanas entre elas. O efeito teste consiste na melhora do desempenho quanto à retenção de informação em indivíduos que são submetidos a provas repetidas sobre determinado assunto (BUTLER, 2010; PASTURA; SANTORO-LOPES, 2013). Foi inicialmente descrito no século XX, principalmente por Gates e Spitzer na década de 1930. No entanto, permaneceu relegado a raros trabalhos de Psicologia Cognitiva até o início deste século. Roediger e Karpicke publicaram revisão sobre o efeito teste na educação médica e demonstraram que ele não se deve à maior frequência de estudo para se fazer várias provas sobre um mesmo tema, mas sim, um efeito intrínseco à memória, devido ao processo ativo para buscar a informação solicitada durante a prova, facilitando a retenção das informações em longo prazo (ROEDIGER; KARPICKE, 2006). Em estudo realizado no laboratório de simulação com 81 alunos de Medicina do sétimo semestre, Kromann, Jensen e Ringsted (2009) demonstraram esse efeito também em habilidades clínicas procedimentais. Os estudantes foram randomizados em dois grupos, o grupo controle fez um curso de atendimento a parada cardíaca com duração de 4 horas e o grupo intervenção fez um curso de 3,5 horas mais 30 minutos de prova. Duas semanas depois, os dois grupos foram submetidos a uma prova prática-oral com o mesmo avaliador e *checklist*. A média das notas do grupo intervenção foi maior em comparação ao grupo controle, 82,8% e 73,3% respectivamente (KROMANN; JENSEN; RINGSTED, 2009). Ainda que os resultados sugiram o efeito teste, não é possível descartar a possibilidade de que os participantes tenham estudado o tema entre as duas fases do estudo.

Mesmo considerando-se o número pequeno de alunos, a prova prática-oral estruturada parece ser comparável a uma estação do OSCE na avaliação de micro-habilidades clínicas de acordo com os resultados observados. Os achados também reforçam a possibilidade de ocorrência de efeito teste em provas práticas procedimentais, conforme descrito por Kromann, Jensen e Ringsted (2009). Novos estudos, com número maior de discentes e com maior número de habilidades, são necessários para confirmar essas conclusões.

### CONSIDERAÇÕES FINAIS

Neste trabalho, a prova prática-oral estruturada com *checklist* apresentou resultados equivalentes a uma estação de OSCE, sugerindo a possibilidade de sua utilização na avaliação de micro-habilidades clínicas em ambiente simulado. Como a prova prática-oral estruturada é de organização mais simples, menor custo, pode ser aplicada por um único professor e tem menor potencial estressante, ela é uma alternativa interessante e confiável para realização de avaliações ao longo do semestre ou ao término de unidades avaliativas. O OSCE, devido à sua complexidade operacional e custos, pode ser utilizado ao final do período letivo, principalmente para a avaliação de competências clínicas.



Outra observação interessante é o possível efeito teste independente do tipo de prova. Considerando a menor complexidade de realização da prova prática-oral estruturada e o efeito teste observado neste estudo, a realização frequente desse tipo de prova associada ao *feedback* imediato pode proporcionar maior aprendizado, transformando a avaliação numa ferramenta de aprendizagem.

## REFERÊNCIAS

AL OMARI, Ahmad; SHAWAGFA, Zuhair M. New experience with objective structured clinical examination in Jordan. *Rawal Medical Journal*, v. 35, n. 1, p. 78-81, jan. 2010.

AL-ELQ, Abdulmohsen H. Medicine and Clinical Skills Laboratories. *Journal of Family & Community Medicine*, v. 14, n. 2, p. 59-63, maio 2007.

AMIRI, Marzieh; NICKBAKHT, Mansoureh. The Objective Structured Clinical Examination: A study on satisfaction of students, faculty members, and tutors. *Life Science Journal*, v. 9, n. 4, p. 4909-4911, dez. 2012.

ANASTAKIS, Dimitri J.; COHEN, Robert; REZNICK, Richard K. The Structured Oral Examination as a Method for Assessing Surgical Residents. *American Journal of Surgery*, New York, v. 162, n. 1, p. 67-70, jul. 1991.

BRAND, Hendrik S.; SCHOONHEIM-KLEIN, Meta. Is the OSCE more stressful? Examination anxiety and its consequences in different assessment methods in dental education. *European Journal of Dental Education*, Oxford, v. 13, n. 3, p. 147-153, set. 2009. DOI: <http://dx.doi.org/10.1111/j.1600-0579.2008.00554.x>.

BUTLER, Andrew C. Repeated Testing Produces Superior Transfer of Learning Relative to Repeated Studying. *Journal of Experimental Psychology: Learning, Memory and Cognition*, Washington D.C., v. 36, n. 5, p. 1118-1133, set. 2010. DOI: <https://doi.org/10.1037/a0019902>.

DAVIS, Margery H. OSCE: The Dundee Experience. *Medical Teacher*, London, v. 25, n. 3, p. 255-261, maio 2003. DOI: <http://dx.doi.org/10.1080/0142159031000100292>.

DAVIS, Margery H.; KARUNATHILAKE, Indika. The Place of the Oral Examination in Today's Assessment Systems. *Medical Teacher*, London, v. 27, n. 4, p. 294-297, jun. 2005. DOI: <http://dx.doi.org/10.1080/01421590500126437>.

ELDER, Andrew. Clinical Skills Assessment in the Twenty First Century. *Medical Clinics of North America*, Philadelphia, v. 102, n. 3, p. 545-558, maio 2018. DOI: <https://doi.org/10.1016/j.mcna.2017.12.014>.

EPSTEIN, Ronald M. Assessment in Medical Education. *New England Journal of Medicine*, Boston, v. 356, n. 4, p. 387-396, jan. 2007.

GORMLEY, Gerry. Summative OSCEs in Undergraduate Medical Education. *Ulster Medical Journal*, Belfast, v. 80, n. 3, p. 127-132, set. 2011.

HARDEN, Ronald M. What Is an OSCE?. *Medical Teacher*, London, v. 10, n. 1, p. 19-22, 1988. DOI: <https://doi.org/10.3109/01421598809019321>.

HARDEN, Ronald M. Misconceptions and the OSCE. *Medical Teacher*, London, v. 37, n. 7, p. 608-610, jul. 2015. DOI: [10.3109/0142159X.2015.1042443](https://doi.org/10.3109/0142159X.2015.1042443).

HARDEN, Ronald M. Revisiting 'Assessment of Clinical Competence Using an Objective Structured Clinical Examination (Osce)'. *Medical Education*, Oxford, v. 50, n. 4, p. 376-379, abr. 2016. DOI: <https://doi.org/10.1111/medu.12801>.

HARDEN, Ronald M.; GLEESON, F. A. Assessment of Clinical Competence Using an Objective Structured Clinical Examination (OSCE). *Medical Education*, Oxford, v. 13, n. 1, p. 41-54, jan. 1979.

HARDEN, Ronald M *et al.* Assessment of Clinical Competence Using Objective Structured Examination. *British Medical Journal*, London, v. 1, n. 5955, p. 447-451, fev. 1975. DOI: <https://dx.doi.org/10.1136%2Fbmj.1.5955.447>.

HASHIM, Rizwan *et al.* Structured Viva as an Assessment Tool: Perceptions of Undergraduate Medical Students. *Pakistan Armed Forces Medical Journal*, v. 65, n. 1, p. 141-144, fev. 2015.

JAYAWICKRAMARAJAH, P. T. Oral Examinations in Medical Education. *Medical Education*, Oxford, v. 19, n. 4, p. 290-293, jul. 1985. DOI: <https://doi.org/10.1111/j.1365-2923.1985.tb01323.x>.

KHAN, Huma Musarrat; MIRZA, Taaha Muddassir. Perceptions of Oral Structured Examination: A Move from Subjectivity to Objectivity. *Pakistan Armed Forces Medical Journal*, v. 67, n. 1, p. 41-46, fev. 2017.

KHAN, Kamran Z. *et al.* The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An Historical and Theoretical Perspective. *Medical Teacher*, London, v. 35, n. 9, p. e1437-1446, set. 2013b. DOI: <https://doi.org/10.3109/0142159x.2013.818634>.

KHAN, Kamran Z. *et al.* The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & Administration. *Medical Teacher*, London, v. 35, n. 9, p. e1447-1463, set. 2013a. DOI: <https://doi.org/10.3109/0142159x.2013.818635>.

KROMANN, Charles B.; JENSEN, Morten L.; RINGSTED, Charlotte. The Effect of Testing on Skills Learning. *Medical Education*, Oxford, v. 43, n. 1, p. 21-27, jan. 2009. DOI: <https://doi.org/10.1111/j.1365-2923.2008.03245.x>.

LEBLANC, Vicki. R.; BANDIERA, Glen W. The effects of examination stress on the performance of emergency medicine residents. *Medical Education*, Oxford, v. 41, n. 6, p. 556-564, jun. 2007. DOI: <https://doi.org/10.1111/j.1365-2923.2007.02765.x>.

LUNENFELD, Eitan *et al.* Assessment of Emergency Medicine: A Comparison of an Experimental Objective Structured Clinical Examination with a Practical Examination. *Medical Education*, Oxford, v. 25, n. 1, p. 38-44, jan. 1991. DOI: <https://doi.org/10.1111/j.1365-2923.1991.tb00024.x>.

MAJUMDER, Md Anwarul A. *et al.* An evaluative study of objective structured clinical examination (OSCE): students and examiners perspectives. *Advances in Medical Education and Practice*, Auckland, v. 10, p. 387-397, jun. 2019. DOI: <https://dx.doi.org/10.2147%2FAMEP.S197275>.

MATHEWS, Lulu; MENON, Janaki; MANI, N. S. Micro-OSCE for assessment of undergraduates. *Indian Pediatrics*, v. 41, p. 159-163, mar. 2004.

MEMON, Muhammed Ashraf; JOUGHIN, Gordon Rowland; MEMON, Breda. Oral Assessment and Postgraduate Medical Examinations: Establishing Conditions for Validity, Reliability and Fairness. *Advances in Health Sciences Education: Theory and Practice*, Boston, v. 15, n. 2, p. 277-289, maio 2010. DOI: <https://doi.org/10.1007/s10459-008-9111-9>.

MILLER, George E. The Assessment of Clinical Skills/Competence/Performance. *Academic Medicine*, Philadelphia, v. 65, n. 9 Supl., p. S63-67, set. 1990. DOI: <https://doi.org/10.1097/00001888-199009000-00045>.

MONDAL, Rakesh *et al.* Comparative analysis between Objective Structured Clinical Examination (OSCE) and Conventional Examination (CE) as a formative evaluation tool in Pediatrics in semester examination for final MBBS students. *Kathmandu University Medical Journal*, Kathmandu, v. 37, n. 1, p. 62-65, jan./mar. 2012. DOI: <https://doi.org/10.3126/kumj.v10i1.6917>.

PASTURA, Patrícia Souza Valle Cardoso; SANTORO-LOPES, Guilherme. O Aprendizado Melhorado Por Provas. *Revista Brasileira de Educação Médica*, Rio de Janeiro, v. 37, n. 3, p. 429-433, set. 2013. DOI: <https://doi.org/10.1590/S0100-55022013000300015>.

PIERRE, Russell *et al.* Student evaluation of an OSCE in paediatrics at the University of the West Indies, Jamaica. *BMC Medical Education*, London, v. 4, p. 22, 2004. DOI: <https://dx.doi.org/10.1186%2F1472-6920-4-22>.

RAZAVI, Seyyed M. *et al.* Station-Based Deconstructed Training Model for Teaching Procedural Skills to Medical Students: A Quasi-Experimental Study. *Advances in Medical Education and Practice*, Auckland, v. 1, p. 17-23, set. 2010. DOI: <https://doi.org/10.2147/amep.s13750>.

ROEDIGER, Henry L.; KARPICKE, Jeffrey D. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, New York, v. 17, n. 3, p. 249-255, mar. 2006. DOI: <https://doi.org/10.1111/j.1467-9280.2006.01693.x>.

SMEE, Sydney. Skill Based Assessment. *British Medical Journal*, London, v. 326, n. 7391, p. 703-706, mar. 2003. DOI: <https://doi.org/10.1136/bmj.326.7391.703>.

THIESSEN, Nils; FISCHER, Martin; HUVENDIEK, Soren. Assessment methods in medical specialist assessments in the DACH region – overview, critical examination and recommendations for further development. *GMS Journal for Medical Education*, Erlangen, v. 36, n. 6, p. 1-23, 2019. DOI: <https://dx.doi.org/10.3205%2Fzma001286>.

TIJANI, Kehinde H. *et al.* A comparison of The Objective Structured Clinical Examination and the Traditional Oral Clinical Examination in a Nigerian University. *Journal of the West African College of Surgeons*, Lagos, v. 7, n. 3, p. 59-72, jul./set. 2017.

TRONCON, Luiz Ernesto. Clinical skills assessment: limitations to the introduction of an "OSCE" (Objective Structured Clinical Examination) in a traditional Brazilian medical school. *São Paulo Medical Journal*, São Paulo, v. 122, n. 1, p. 12-17, 2004. DOI: <https://doi.org/10.1590/S1516-31802004000100004>.

VAN DER VLEUTEN, Cees P. *et al.* The Assessment of Professional Competence: Building Blocks for Theory Development. *Best Practice & Research: Clinical Obstetrics & Gynaecology*, Amsterdam, v. 24, n. 6, p. 703-719, dez. 2010. DOI: <https://doi.org/10.1016/j.bpobgyn.2010.04.001>.

WASS, Val *et al.* Assessment of Clinical Competence. *Lancet*, London, v. 357, n. 9260, p. 945-949, mar. 2001. DOI: [https://doi.org/10.1016/s0140-6736\(00\)04221-5](https://doi.org/10.1016/s0140-6736(00)04221-5).

WASS, Val *et al.* Achieving Acceptable Reliability in Oral Examinations: An Analysis of the Royal College of General Practitioners Membership Examination's Oral Component. *Medical Education*, Oxford, v. 37, n. 2, p. 126-131, fev. 2003. DOI: <https://doi.org/10.1046/j.1365-2923.2003.01417.x>.

YOUNG, Ian *et al.* The benefits of a peer-assisted mock OSCE. *The Clinical Teacher*, Oxford, v. 11, n. 3, p. 214-218, jun. 2014. DOI: <https://doi.org/10.1111/tct.12112>.

A prova prática-oral estruturada é comparável ao exame clínico objetivo estruturado na avaliação de micro-habilidades clínicas?

*Flávia Soares de Matos, Antonio Toledo Jr.*

---

**Flávia Soares de Matos**

Médica Pediatra. Mestre em Ensino em Saúde pela Universidade José do Rosário Vellano em Belo Horizonte (UNIFENAS/BH). Professora do Curso de Medicina da UNIFENAS em Belo Horizonte. Coordenadora do Laboratório de Simulação. Professora colaboradora do Mestrado Profissional em Ensino em Saúde – UNIFENAS/BH.

flavia.matos@unifenas.br

**Antonio Toledo Jr.**

Médico Infectologista. Mestre em Saúde Pública/Epidemiologia – UFMG (1998). Doutor em Medicina Tropical e Infectologia – UFTM (2008). Professor TI do Curso de Medicina da Universidade José do Rosário Vellano em Belo Horizonte (UNIFENAS/BH). Coordenador do Internato de Clínica Médica e do Programa de Pós-Graduação em Ensino em Saúde. Linhas de pesquisa: avaliação discente e de programas e inteligência emocional.

antonio.toledo@unifenas.br