

Avaliação do estudante de graduação nas profissões da saúde: mudanças associadas ao desenvolvimento docente

Undergraduate student assessment in the health professions: changes associated to faculty development

Rodrigo Humberto Flauzino^{1,2} 

roflauzino@usp.br

Maria Paula Panúncio-Pinto^{1,3} 

mapaula@fmrp.usp.br

Valdes Roberto Bollela^{1,4} 

vbollela@fmrp.usp.br

Luiz Ernesto de Almeida Troncon^{1,4} 

ledatron@fmrp.usp.br

RESUMO

Introdução: A avaliação dos estudantes nos cursos de graduação nas profissões da saúde tem importância essencial, podendo cumprir as finalidades somativa, formativa e informativa. No entanto, há poucos dados disponíveis sobre como ela se dá nas instituições brasileiras. Uma ação de desenvolvimento docente (DD), que consistiu em ciclos de oficinas para gestores e professores sobre programas institucionais de avaliação, com ênfase formativa, proporcionou oportunidade para explorar esse panorama.

Objetivo: Este estudo teve como objetivo caracterizar as práticas avaliativas em cursos de graduação em saúde no Brasil, particularizando responsabilidades institucionais, métodos de avaliação e *feedback* ao estudante, e identificando mudanças nesses aspectos associadas às atividades de DD.

Método: Trinta cursos de graduação em saúde de nove diferentes instituições de ensino superior (Regiões Nordeste, Sudeste e Sul) participaram dessa iniciativa, cujos gestores responderam a dois questionários sobre os processos avaliativos desenvolvidos. O primeiro questionário, aplicado antes das oficinas de DD, gerou 23 diagnósticos situacionais acerca do contexto, das responsabilidades sobre a avaliação, dos métodos adotados e do domínio de práticas de *feedback* fornecido aos estudantes. O segundo, aplicado um ano após o último ciclo de capacitação, gerou 36 conjuntos de respostas, expressando a percepção sobre mudanças. As respostas ao primeiro instrumento foram analisadas por meio da técnica de análise de conteúdo de Bardin. As do segundo foram expressas de modo quantitativo.

Resultado: Antes das oficinas, em 70% dos cursos as práticas avaliativas eram de responsabilidade dos docentes, com pouca participação institucional. Mais da metade dos cursos focava a avaliação somativa, com limitações na formativa e no *feedback*. Os métodos de avaliação de conhecimentos e de habilidades clínicas eram variados, enquanto os do domínio afetivo se mostravam imprecisos. Um ano após as ações de DD, notou-se aumento expressivo da participação institucional, com maior equilíbrio nas responsabilidades entre professores e instâncias centrais. Na percepção de 89% dos gestores, o domínio das técnicas de *feedback* melhorou. Ademais, os diferentes métodos avaliativos tornaram-se mais adequados para cada fim.

Conclusão: Aspectos da avaliação, como reduzida responsabilidade institucional, limitada prática do *feedback* fornecido ao estudante e adequação dos métodos, sofreram mudanças expressivas que, na percepção dos gestores, foram associadas à participação de professores e gestores nas atividades de DD.

Palavras-chave: Avaliação Educacional; Estudantes de Ciências da Saúde; Ensino Superior; Profissões de Saúde; Capacitação de Professores.

ABSTRACT

Introduction: Student assessment in undergraduate courses in the health professions holds essential importance, ideally fulfilling summative, formative and informative purposes. However, data available on how it occurs in Brazilian institutions are scarce. A Faculty Development (FD) action, which consisted of cycles of workshops for course managers and teachers focusing on institutional assessment programs, with a formative emphasis, provided an opportunity to explore these settings.

Objective: To characterize the assessment practices in undergraduate health courses in Brazil, detailing institutional responsibilities, assessment methods, and student feedback, and identifying changes in these aspects associated with FD activities.

Method: Thirty undergraduate health courses from nine different higher education institutions (Northeast, Southeast and South regions) participated in this initiative and their managers responded to two questionnaires about the assessment processes in place. The first questionnaire, applied before the FD workshops, generated 23 situational diagnoses about context, assessment responsibilities, methods adopted and mastery of techniques for feedback to students. The second, applied one year after the last FD activity, generated 36 sets of responses expressing perceptions on changes. Responses to the first instrument were analyzed using Bardin's Content Analysis technique. Data from the second were analyzed quantitatively.

Result: Before the workshops, in 70% of the courses, student assessment were the responsibility of teachers, with little institutional participation. More than half of the courses focused on summative assessment, with limitations on formative assessment and feedback to students. Methods for assessing knowledge and clinical skills were quite varied, while those for the affective domain were imprecise. One year after the FD activities, there was a considerable significant increase in institutional participation, with a greater balance in responsibilities between teachers and central instances. In the perception of 89% of managers, mastery of feedback techniques improved. Furthermore, the different assessment methods have become more appropriate for each purpose.

Conclusion: Aspects of assessment, such as reduced institutional responsibility, limited practice of feedback to students and adequacy of methods, underwent significant changes, which, in the perception of administrators, were associated with the participation of faculty and managers in FD activities.

Keywords: Educational measurement; Health Occupations Student; Higher Education; Health professions; Professional Training.

¹ Centro de Desenvolvimento Docente para o Ensino, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil (FMRP-USP).

² Centro de Apoio Educacional e Psicológico, Faculdade de Medicina de Ribeirão Preto (FMRP-USP).

³ Departamento de Ciências da Saúde, Faculdade de Medicina de Ribeirão Preto (FMRP-USP).

⁴ Departamento de Clínica Médica, Faculdade de Medicina de Ribeirão Preto (FMRP-USP).

Editora-chefe: Rosiane Viana Zuza Diniz. | Editor associado: Maria Helena Senger.

Recebido em 20/08/24; Aceito em 07/04/25. | Avaliado pelo processo de double blind review.

INTRODUÇÃO

Atualmente, a avaliação do estudante é entendida como um processo de colheita, análise e interpretação de informações sobre o desempenho dele nos domínios cognitivo, psicomotor e afetivo-attitudinal, bem como nas competências relacionadas a esses aspectos. Ela visa cumprir finalidades específicas, entre as quais se destacam: a tomada de decisão sobre o progresso do estudante (avaliação somativa), o fomento à aprendizagem contínua (avaliação formativa) e o aperfeiçoamento do processo educacional (avaliação informativa)^{1,2}.

Nos cursos brasileiros de graduação nas profissões de saúde, ainda predominam práticas marcadas por um paradigma tradicional, em que a avaliação é de responsabilidade exclusiva do professor da unidade curricular, e prevalecem os exames somativos, com foco quase exclusivo no domínio cognitivo e realizados no âmbito das disciplinas e dos estágios profissionalizantes³.

Apesar das iniciativas para aprimorar os processos avaliativos, como a expansão e a consolidação do Teste de Progresso (TP) nos cursos de Medicina⁴, ou das contribuições do ensino híbrido durante a pandemia, que ampliaram as possibilidades de realizar avaliação formativa mediada por tecnologias, promovendo novas interações professor-estudante⁵⁻⁷, ainda há muito a avançar na realidade brasileira.

No paradigma antes mencionado³, a avaliação de habilidades clínicas, procedimentais e atitudinais, apesar de sua grande importância, ocorre de maneira esporádica e menos sistematizada, sendo pouco frequente o uso de métodos estruturados adequados para esses fins. A avaliação formativa costuma ser escassa, e não há investimento institucional suficiente para que os docentes dominem as técnicas de provisão de *feedback* fornecido ao estudante.

Como nesse modelo a avaliação é responsabilidade do professor ou da coordenação da unidade curricular, faltam instâncias institucionais que coordenem os processos, integrem as informações e decidam sobre o progresso do estudante e o aprimoramento das atividades curriculares². As avaliações geralmente resultam em um escore (nota ou média de notas), que determina a aprovação ou a reprovação do graduando, conforme o ponto de corte definido pela instituição de ensino em seu estatuto.

Assim sendo, as práticas avaliativas encontram-se distantes das principais diretrizes nacionais e internacionais sobre o tema. No cenário internacional, a International Association for Health Professions Education (AMEE) propõe critérios para reconhecer práticas inovadoras por meio do *Aspire Awards*⁸. Já a World Federation for Medical Education (WFME) periodicamente disponibiliza recomendações específicas para os processos de ensino, aprendizagem

e avaliação⁹. Essas organizações incentivam a criação de programas ou sistemas institucionais de avaliação que permitam o cumprimento de propósitos específicos e a valorização da avaliação formativa, aquela que promove a aprendizagem (*assessment for learning*)^{3,10,11}.

Essas recomendações indicam que os programas de avaliação devem apoiar a missão das instituições na educação nas profissões da saúde, com o intuito de melhorar a saúde das pessoas. Além disso, é necessário que aprimorem e criem oportunidades de aprendizagem, garantindo a competência dos estudantes conforme progridem, com rigoroso e contínuo controle de qualidade. Esse compromisso com o conhecimento e a inovação inclui a divulgação das boas práticas⁸ por meio de programas robustos de avaliação capazes de assegurar, impulsionar, orientar e otimizar os processos educacionais⁹.

No Brasil, as Diretrizes Curriculares Nacionais (DCN) para os cursos de graduação em Medicina¹² enfatizam a necessidade de monitoramento e de avaliação contínuos do curso e dos estudantes, de forma progressiva, nos aspectos cognitivos, psicomotores e afetivo-attitudinais, estabelecendo a obrigatoriedade de oferta e manutenção “permanente de Programa de Formação e Desenvolvimento da Docência em Saúde”.

A atualização das DCN para outros cursos de graduação nas profissões da saúde segue a mesma direção. O Conselho Nacional de Saúde (CNS) publicou recomendações que destacam a “avaliação com caráter processual e formativo”¹³, o uso de metodologias ativas e a definição de critérios para o acompanhamento e a avaliação dos processos de ensino-aprendizagem. O CNS recomenda também a manutenção de atividades de desenvolvimento docente (DD), e, nesse sentido, deve-se destacar que seria igualmente importante que os programas de pós-graduação, responsáveis pela formação de mestres e doutores, os futuros professores e gestores, dediquem-se à discussão e reflexão sobre a formação em saúde.

No entanto, são escassos os dados objetivos sobre como se dá, atualmente, a avaliação do estudante nos diversos cursos de graduação na área da saúde. Não há evidências claras acerca das características das práticas avaliativas e das iniciativas de aprimoramento adotadas para incorporar as recomendações expressas.

Nesse contexto, o DD desempenha papel fundamental para contribuir para a mudança do paradigma de avaliação dos estudantes nos cursos de graduação em saúde no Brasil. Essa transformação só será possível com oportunidades reais e regulares de capacitação dos professores nessa área.

Uma iniciativa de DD que envolveu professores e gestores de diversas instituições em diferentes estados

brasileiros possibilitou obter dados para elaborar um panorama situacional das práticas avaliativas em um número considerável de cursos de graduação na área da saúde. Essa ação permitiu identificar sinais de mudanças tanto na manutenção e consolidação do que já funcionava bem quanto na melhoria de aspectos, constituindo assim um movimento de ruptura do padrão tradicional.

Portanto, o objetivo deste artigo foi verificar a influência de um projeto de DD nas práticas avaliativas em cursos de graduação nas profissões da saúde no Brasil, de modo a caracterizar aspectos como as responsabilidades institucionais, os métodos de avaliação e as técnicas de *feedback*, e identificar mudanças associadas à participação em atividades de DD a partir da visão dos gestores dos cursos.

MÉTODOMetodologiaContexto e desenho

O Centro de Desenvolvimento Docente para o Ensino (CDDE) da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (FMRP-USP) criou uma linha de trabalho focada na capacitação de professores, com ênfase na avaliação do estudante e destaque à avaliação formativa¹⁴.

Em 2018, iniciou-se uma nova frente de atuação específica, destinada ao preparo de professores e gestores de cursos para a atuação em avaliação programática institucional^{2,3,10,11}. Essa ação fez parte de um projeto contemplado com recursos da iniciativa Latin American Grants (LAG) do National Board of Medical Examiners (NBME) dos Estados Unidos. Os fundos permitiram realizar atividades de DD, disseminando essas reflexões e práticas em contextos brasileiros da educação em saúde, por meio de oficinas (manuscrito em preparação) que contaram com participantes, incluindo professores e gestores de cursos de graduação.

Esses participantes eram provenientes de nove diferentes instituições de ensino superior (IES), localizadas: uma da Região Nordeste, uma da Região Sul e as demais do Sudeste. Dessas instituições, sete eram públicas (estaduais ou federais) e duas privadas (sendo uma de natureza confessional). Esses profissionais atuavam em 30 cursos de graduação de 13 áreas da saúde distintas, incluindo: Ciências Biomédicas, Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Gerontologia, Informática Biomédica, Medicina, Medicina Veterinária, Nutrição, Odontologia, Terapia Ocupacional e Serviço Social.

As oficinas, realizadas entre novembro de 2018 e janeiro de 2020, seguiram um programa de conteúdos centrados na avaliação do estudante, utilizando estratégias de ensino-aprendizagem ativas e participativas (Quadro 1). Ao todo, ocorreram 11 ciclos de oficinas presenciais, inicialmente

na sede da instituição proponente e, posteriormente, nas escolas parceiras. Cada ciclo teve duração de dois dias, com carga horária total de 20 horas. Os objetivos gerais dessa ação incluíram a formação de massa crítica de professores e gestores dos cursos de graduação nas profissões da área da saúde, capacitando-os para o planejamento de projetos relacionados à avaliação programática dos estudantes em seus cenários educacionais. Isso visou promover melhor compreensão dos conceitos e das finalidades da avaliação, além do aprimoramento de práticas já existentes ou da introdução de novas abordagens.

Participantes

As ações de DD mobilizaram 359 participantes dos vários cursos das profissões da saúde. No entanto, buscando a caracterização de cunho institucional da avaliação do estudante, este estudo focou as visões dos gestores sobre os processos avaliativos do discentes em seus cursos.

Quadro 1. Conteúdo e estratégias das oficinas de DD sobre avaliação do estudante, no âmbito do Projeto LAG-NBME, desenvolvido pelo CDDE-FMRP-USP.

Conteúdo
Conceitos de avaliação programática e sistemas de avaliação
Conceitos-chave sobre avaliação educacional, critérios de qualidade na avaliação do estudante e escolha de métodos
Alinhamento entre o currículo e a avaliação do estudante
Avaliação de habilidades cognitivas
Avaliação estruturada de competências clínicas
Avaliação nos cenários da prática profissional
Avaliação formativa e técnicas de <i>feedback</i>
Portfólio e tutoria acadêmica (registro de dados da avaliação do estudante)
Experiências de sucesso rumo à avaliação programática
O futuro da avaliação programática institucional: onde estamos e onde podemos chegar.
Estratégias visando à aprendizagem
Leitura de textos
Preleções dialogadas
Discussões em grupos
Exercícios individuais e em grupos
Protagonização (<i>role playing</i>)
Discussões plenárias
Estratégias de avaliação
Formulários de caracterização da avaliação do estudante nos cursos de graduação dos participantes, de autoavaliação inicial e final da aprendizagem, e de apreciação da qualidade das atividades e da oficina.

Fonte: Elaborado pelos autores.

Para isso, antes das oficinas de DD, os gestores responderam a um questionário estruturado com perguntas abertas sobre os aspectos que permitiriam caracterizar como eram, naquele momento, as condições da avaliação do estudante em seus cursos ("diagnóstico situacional inicial"). Dos 30 cursos participantes, foram obtidos 23 diagnósticos nessa fase inicial, correspondendo então a 76,6% dos cursos participantes.

Em um segundo momento, um ano após a finalização do último ciclo de oficinas, os gestores desses mesmos 30 cursos foram convidados a responder um novo questionário, dessa vez eletrônico, para avaliar especificamente as mudanças nas práticas de avaliação dos estudantes, que poderiam ser resultantes ou não da participação dos gestores e docentes dos diferentes cursos nas oficinas de DD. Essa etapa do estudo ocorreu entre fevereiro e maio de 2020.

Nessa fase, foram obtidos 36 questionários respondidos, pois, em seis dos 30 cursos participantes, tanto o gestor original, que havia respondido ao instrumento inicial, quanto o atual coordenador responderam independentemente ao instrumento. Destaque-se que, nesse segundo momento, foram obtidos dados de todos os 30 cursos, dos quais seis contaram com mais de uma percepção. Nos seis cursos com duas respostas de gestores, ambas foram analisadas, e não se notaram grandes divergências. Portanto, as respostas foram consideradas independentes para proporcionar uma visão mais abrangente do curso.

Instrumentos

• *Questionário 1*: Instrumento de diagnóstico situacional inicial, composto por questões abertas, a serem respondidas discursivamente, em papel, pelos gestores de cursos. As perguntas foram divididas em quatro seções: caracterização geral do curso, caracterização geral da avaliação, métodos e instrumentos predominantemente usados, e suporte de DD. Para este relato, selecionaram-se quatro questões específicas que oferecem uma visão das condições de avaliação do estudante:

1) A avaliação do estudante é da responsabilidade das unidades curriculares (módulos, disciplinas, estágios) ou da instituição como um todo (coordenação do curso ou direção da escola)?

2) Descreva os principais métodos utilizados no curso para a avaliação cognitiva, para a avaliação de habilidades procedimentais e clínicas, e para a avaliação do comportamento profissional.

3) Há equilíbrio entre as finalidades somativa (aprovar/reprovar/remediação ou recuperação) e formativa?

4) Há domínio satisfatório de técnicas de *feedback* e devolutiva ao estudante por parte dos membros do corpo docente e dos preceptores? Descreva e comente.

• *Questionário 2*: Instrumento eletrônico (Google Forms) para caracterizar as condições de avaliação do estudante após as atividades de DD. Ele foi aplicado um ano após o último ciclo de oficinas e manteve as mesmas seções e temas do questionário 1, mas utilizou questões objetivas de múltipla escolha, visando captar as percepções dos efeitos das oficinas de DD nos cursos. Cada pergunta permitia apenas uma resposta. Para este manuscrito, foram selecionadas questões específicas para identificar possíveis mudanças e avanços nos processos avaliativos:

1) Em relação à avaliação do estudante, a responsabilidade na sua instituição é ou continua sendo (responda escolhendo somente uma alternativa):

- Totalmente das unidades curriculares (módulos, disciplinas, estágios).
- Predominantemente das unidades curriculares, mas com a participação de instâncias da instituição (coordenação de curso ou direção da escola).
- Equilibrada entre as unidades curriculares e a instituição.
- Totalmente da instituição (coordenação do curso ou direção da escola).
- Outra situação.

2) Houve avanços em relação ao domínio de técnicas de *feedback* e devolutiva ao estudante por parte dos membros do corpo docente e de preceptores na sua instituição, de modo geral, após as atividades de DD?

3) Houve avanços em relação ao equilíbrio entre as finalidades *somativa* e *formativa* da avaliação do estudante, após as atividades de DD?

4) Houve avanços quanto aos métodos utilizados no curso para a avaliação cognitiva, após as atividades de DD?

5) Houve avanços quanto aos métodos utilizados no curso para a avaliação de habilidades procedimentais e clínicas, após as atividades de DD?

6) Houve avanços quanto aos métodos utilizados no curso para a avaliação do comportamento profissional, após as atividades de DD?

As alternativas de resposta a essas questões anteriores eram:

- Não houve avanços, persistindo necessidade de adequação.
- Não houve avanços, pois o que havia já estava adequado.
- Houve avanços, independentemente da participação nas atividades de DD.
- Houve avanços, em decorrência da participação nas atividades de DD.
- Outra situação.

Antes de serem enviados aos gestores, os instrumentos foram revisados por sete juízes (professores), incluindo dois especialistas em ciências biomédicas básicas e cinco que atuavam na etapa clínica de diferentes cursos de saúde. Eles avaliaram a apresentação, o formato, a pertinência, a clareza e o conteúdo das perguntas, contribuindo com seus conhecimentos e experiência. Suas observações permitiram revisões que asseguraram que os questionários atendessem aos requisitos de validade de face e de conteúdo^{15,16}.

Análise dos dados

- **Questionário 1:** Os documentos originais impressos contendo os 23 questionários respondidos tiveram os seus conteúdos copiados para uma planilha Excel (Microsoft Office 365, versão 2016).

Na análise temática de conteúdo, adotou-se a estratégia de análise qualitativa, de acordo com a perspectiva de Bardin¹⁷, que consiste em:

Um conjunto de técnicas de análise das comunicações visando obter, por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) dessas mensagens (p.44).

Na aplicação dessa técnica de análise, após a leitura inicial do material ("exploração flutuante"), outras abordagens mais detalhadas e aprofundadas foram realizadas, com a identificação dos temas e a "contagem de vários temas ou itens de significação"¹⁷. É importante considerar que, nesse caso, esses temas ou categorias foram dados *a priori* pelos questionários. A identificação da frequência de termos surgidos explicitou se a "responsabilidade pela avaliação do estudante" era mais da alçada dos professores e coordenadores de unidades curriculares ou era compartilhada com outras instâncias do curso ou da instituição, como comitês de gestão ou de avaliação. Esse mesmo procedimento foi aplicado às perguntas sobre o "equilíbrio entre as finalidades somativa e formativa da avaliação" e o "domínio satisfatório das técnicas de *feedback*".

Para ilustrar os temas, selecionaram-se recortes discursivos específicos (transcrição literal), seguidos por um código: letras iniciais que designam os cursos e um número atribuído durante a análise.

No caso das perguntas sobre os métodos de avaliação, os resultados serão apresentados em tabelas separadas para os domínios cognitivo e de habilidades procedimentais ou clínicas, contendo as expressões agrupadas e contabilizadas, que demonstram as estratégias predominantes nos cursos. Para o domínio do comportamento profissional, elaborou-

se um quadro contendo as estratégias mencionadas pelos participantes, uma vez que, devido à diversidade de formatos mencionados, não foi possível identificar a frequência precisa de cada uma. Assim, diante do universo de respostas analisadas, em cada tabela ou quadro, apresentaram-se as frequências dos métodos declarados.

- **Questionário 2:** Os dados do segundo questionário, aplicado eletronicamente, foram extraídos e tabulados em uma planilha Excel (Microsoft Office 365, versão 2016). As respostas às perguntas explicitadas anteriormente foram analisadas, e os resultados serão apresentados em porcentagens do total de respondentes.

Aspectos éticos

O projeto de pesquisa foi submetido à apreciação do Comitê de Ética em Pesquisa (CEP) da instituição, sendo aprovado em 6 de março de 2020 (Processo nº 3902020; CAAE nº 29464920.6.0000.5440).

Considerando que parte dos dados já havia sido colhida quando das oficinas já realizadas (para a finalidade exclusiva de aprimorar as atividades de DD), a aprovação deste projeto inclui a autorização pelo CEP de dispensa de assinatura de Termo de Consentimento Livre e Esclarecido (TCLE) aos participantes que responderam ao primeiro questionário.

Quanto ao segundo questionário, os participantes respondentes foram devidamente informados sobre o caráter e os objetivos do estudo, e convidados a participar, expressando a sua concordância mediante aceite eletrônico do TCLE específico, elaborado conforme Resolução nº 510, de 7 de abril de 2016, do CNS, relativa à pesquisa com seres humanos, e previamente submetido ao CEP e aprovado.

Todos os procedimentos de pesquisa foram realizados considerando os riscos, desconfortos e benefícios envolvidos. Dado o tema sensível sobre avaliação da aprendizagem dos estudantes, respeitaram-se a hierarquia e a dependência dos profissionais em relação às suas instituições, assegurando também anonimato e confidencialidade no tratamento das informações obtidas.

RESULTADOS

Responsabilidade pela avaliação do estudante

Dentre o conjunto de respostas de caracterização inicial (23 cursos), em 16 (70%) foi mencionado que a avaliação do estudante era totalmente de responsabilidade do professor e da coordenação das unidades curriculares, como ilustra o seguinte excerto: "A avaliação é exclusiva do docente da disciplina" (TO19). Em sete cursos (30%), as respostas indicaram iniciativas partilhadas, com a criação de comitês, grupos de trabalho ou comissões institucionais

para esse fim, como ilustra este comentário: *“A avaliação dos estudantes é de responsabilidades de todas as instâncias envolvidas na formação ao longo de todo o curso”* (NU6).

As respostas ao segundo questionário mostraram mudanças nos cursos após as ações de DD. A responsabilidade exclusiva dos professores ou das coordenações das unidades curriculares diminuiu de 70% para 45% (17/36 respostas), enquanto o compartilhamento com a instituição aumentou de 30% para 53% (19/36 respostas).

Equilíbrio entre as finalidades somativa e formativa da avaliação do estudante

Dentre o conjunto de respostas ao primeiro questionário, foi possível identificar que em 12 (52,2%) não se percebia equilíbrio, predominando a finalidade somativa, como mostra este comentário: *“Em algumas disciplinas são realizadas avaliações somativas, incluindo discussões após atividades em grupos e avaliações sem atribuição de notas, entretanto ainda não há equilíbrio dessas avaliações formativas e somativas em todas as disciplinas”* (F13).

Seis (26,1%) mostraram haver esse balanceamento entre as duas finalidades, como indica o seguinte trecho: *“A avaliação somativa e formativa é realizada em todas as disciplinas profissionalizantes, utilizando o instrumento escolhido pelos docentes”* (EN23).

Já no restante (21,7%), foram apresentadas descrições genéricas que não permitiram uma conclusão definida, uma vez que alguns discursos dos gestores apontam para a necessidade de estudos sobre esse aspecto em seus cursos, como em: *“Uma vez mais, minha resposta é apenas ilustrativa (parte de um todo, que não consigo detalhar)”* (ME21).

Por sua vez, como destaque, as respostas ao segundo questionário mostraram um aumento de 52,2% para 70% (25/36) no reconhecimento de equilíbrio entre as finalidades somativa e formativa da avaliação do estudante, depois das ações de DD, na perspectiva dos gestores.

Domínio das técnicas de feedback e devolutiva ao estudante

Do conjunto de respostas ao questionário de diagnóstico inicial, em 12 (52,2%) as respostas indicaram claramente a ausência de domínio das técnicas de *feedback* por parte dos docentes, como mostra este comentário: *“Não aplicamos essas técnicas de avaliação”* (IN5).

Seis (26,1%) mostraram haver domínio, como em: *“Muitos docentes se capacitaram e atualmente apresentam esse perfil”* (ME1).

Em cinco (21,7%), houve respostas gerais, ora indicando a necessidade de melhorias, ora reconhecendo não saber como

as devolutivas se manifestam nos cursos, ou ainda indicando heterogeneidade: *“no entanto, não temos como responder esta questão de forma qualificada. Tem-se mostrado um desafio conseguir acessar, em detalhes, as metodologias de avaliação utilizadas pelos docentes”* (TO17).

Já as respostas do segundo questionário mostraram considerável avanço no domínio das práticas de *feedback* depois das ações de DD, na perspectiva de 89% dos gestores (32/36 respostas).

Métodos utilizados para a avaliação no domínio cognitivo

Nas respostas ao questionário de caracterização inicial produzidas pelos gestores, houve relatos de vários métodos que costumavam ser utilizados para a avaliação do estudante no domínio cognitivo. Dentre as menções discursivas, emergiu a utilização de testes de múltipla escolha e de provas com questões abertas (Tabela 1). A partir da análise, foram contabilizados os termos de mesmo sentido sobre métodos, o que gerou um total de 77 registros (100,0%).

Com a aplicação do segundo questionário, depois das ações de DD, na apreciação de 75% (27/36) dos gestores, foram percebidos avanços na escolha e no uso de métodos de avaliação para o âmbito cognitivo.

Tabela 1. Frequência de registros sobre os métodos utilizados para avaliar o estudante no domínio cognitivo em 23 cursos de graduação na área das ciências da saúde, cujos gestores participaram das oficinas de DD. Os dados são apresentados em números absolutos (N) e, entre parênteses, em porcentagens do total de menções.

MÉTODOS UTILIZADOS NOS DIFERENTES CURSOS DE GRADUAÇÃO	
Avaliação do domínio cognitivo	N (%)
Testes de múltipla escolha	22 (28,6)
Provas com questões abertas (curtas, médias, longas)	21 (27,3)
Projetos experimentais e relatórios (científicos, de estágio, informativos, relatos, manifestações significativas, pesquisas de campo)	12 (15,6)
Seminários (apresentações orais, individuais ou em grupo)	7 (9,1)
Casos clínicos (reais, interpretação de dados)	7 (9,1)
PBL, TBL	4 (5,2)
Outros (Teste de Progresso, portfólios, jogos, revisões de literatura)	4 (5,2)
TOTAL DE REGISTROS	77 (100,0)

PBL: *problem-based learning*; TBL: *team-based learning*.
Fonte: Elaborada pelos autores.

Métodos utilizados para a avaliação de habilidades procedimentais e clínicas

As respostas dos gestores indicaram a utilização de muitas estratégias para avaliar as habilidades procedimentais e clínicas (Tabela 2), predominando as provas práticas e de laboratório e o Exame Objetivo Estruturado de Habilidades Clínicas (*Objective Structured Clinical Examination* – OSCE), este, principalmente, nos cursos de Medicina. Algumas narrativas dos gestores apontam que os métodos usados para avaliação desse domínio carecem de coerência (como “avaliação cognitiva”; TP), uma vez que parece não existir a observação do estudante realizando os procedimentos propriamente ditos. Todavia, elas foram analisadas tal como surgiram nas escritas. Assim, consideraram-se as diferentes menções dos 23 formulários analisados, gerando 59 unidades de registros (100,0%).

Já na aplicação do segundo questionário, os dados obtidos indicaram que houve mudanças positivas quanto aos métodos utilizados pelos cursos para a avaliação de habilidades procedimentais e clínicas. Para 55,6% (20/36) dos gestores, esses avanços foram decorrentes diretos da participação das suas equipes docentes nas atividades de DD.

Métodos utilizados para a avaliação do comportamento profissional

As respostas ao questionário inicial revelaram heterogeneidade na descrição dos métodos e instrumentos utilizados na avaliação do estudante, dificultando a quantificação de dados. Alguns deles são estratégias conhecidas na educação das profissões da saúde, enquanto outros foram customizados pelos cursos para fases específicas, conforme os objetivos educacionais (Quadro 2).

No segundo questionário, 58,30% (21/36) dos gestores registraram que emergiram avanços quanto à utilização de

métodos nos cursos para a avaliação do comportamento profissional. Para 11,40% (4/36), não houve avanços, persistindo a necessidade de adequações.

Tabela 2. Frequência de registros sobre os métodos utilizados para avaliar o estudante no domínio procedimental e de habilidades clínicas em 23 cursos de graduação na área das ciências da saúde, cujos gestores participaram das oficinas de DD. Os dados são apresentados em números absolutos (N) e, entre parênteses, em porcentagens do total de menções.

MÉTODOS UTILIZADOS NOS DIFERENTES CURSOS	
Avaliação das habilidades procedimentais e clínicas	N (%)
OSCE	13 (22,0)
Provas práticas e práticas de laboratório	13 (22,0)
Simulações	8 (13,6)
Relatórios, fichas de registro, acompanhamento da evolução do paciente	5 (8,5)
Casos clínicos	4 (6,8)
Provas escritas, objetivas	4 (6,8)
Observação dos estudantes em atendimento ao paciente, DOPS, Mini-CEX	3 (5,1)
Avaliação cognitiva (prova teste ou dissertativa)	2 (3,4)
Portfólio	2 (3,4)
Trabalhos coletivos e seminários	2 (3,4)
Outros (Teste de Progresso, avaliação do produto, produção de TCC)	3 (5,1)
TOTAL DE REGISTROS	59 (100,0)

OSCE: *Objective Structured Clinical Examination*; DOPS: *Direct Observation of Procedural Skills*; Mini-CEX: *Mini-Clinical Evaluation Exercise*; TCC: trabalho de conclusão de curso.
Fonte: Elaborada pelos autores.

Quadro 2. Métodos utilizados para avaliar o estudante no domínio atitudinal e do comportamento profissional em 23 cursos de graduação da área das ciências da saúde, cujos gestores participaram das oficinas de DD. Indicam-se as denominações dos métodos ou instrumentos e os descritores apresentados pelos gestores.

Método ou instrumento utilizado	Menções às atitudes e aos comportamentos avaliados
Avaliações de 360 graus	“Avaliação do estudante por todos à sua volta, como superiores, subordinados, funcionários, pacientes e pelo próprio estudante avaliado e seus pares” (ME1).
Conselhos de classe, que atribuem notas nos estágios	“Avaliação do comportamento profissional do estudante, presenças” (OD13).
Disciplina do eixo de ciências humanas e tutoria acadêmica	“Questões relacionadas a ética profissional, pontualidade, comportamento atitudinal trabalhadas nesses contextos” (CI2).
Estágios profissionalizantes, com registros em formulários desenvolvidos e adaptados para cada especificidade	“Comportamento profissional. A avaliação é subjetiva e realizada pelos preceptores nos serviços onde há as atividades” (NU6). “Pontualidade, trabalho em equipe interprofissional, comportamento profissional, organização do ambiente de trabalho, proatividade, relação com colegas, equipe e supervisores, conduta clínica. Interesse na aprendizagem, assiduidade e pontualidade” (FI20).

Continua...

Quadro 2. Continuação.

Método ou instrumento utilizado	Menções às atitudes e aos comportamentos avaliados
Ficha de avaliação individual estruturada	<i>"Pontualidade, assiduidade, participação nas atividades, relacionamento (colegas, pacientes, preceptores, professores), motivação, ética nas relações pessoais e no desenvolvimento, organização na execução das atividades"</i> (FA11). <i>"Aspectos atitudinais (eventualmente, contam com apoio do setor de psicologia)"</i> (EN8). <i>"Alguns estágios utilizam ficha com parâmetros preestabelecidos para avaliação do comportamento profissional"</i> (ME2).
Instrumentos próprios da disciplina	<i>"Consideram aspectos comportamentais, atitudinais, habilidades, ética e aplicação de conhecimentos"</i> (EN23).
Mini-CEX, uso no ciclo clínico e internato	<i>"Habilidades clínicas e comportamento do estudante frente ao paciente"</i> (ME1).
Registros de prática e de reflexão (aspectos observados em <i>checklist</i> que quantifica construtos, com <i>feedback</i> , com fim formativo ou somativo)	<i>"Postura ética, relacional (com equipe e profissionais), capacidade de planejar, saber fazer, desempenhar, analisar, utilizar referencial teórico prático para realização da prática"</i> (TO19).
Relatórios analíticos; manifestações significativas; participação em trabalhos coletivos (seminários e outros)	<i>"Uso correto de EPI, pontualidade do início e fim do atendimento, frequência, comportamento"</i> (OD13).
Supervisão docente para o TCC	<i>"Acompanhamento do desempenho diário do estudante no laboratório, no qual o supervisor avalia o seu comportamento quanto à pontualidade, à atuação em equipe, ao cumprimento de compromissos acordados e à ética profissional"</i> (CI2).

EPI: equipamento de proteção individual; Mini-CEX: *Mini Clinical Evaluation Exercise*; TCC: trabalho de conclusão de curso.
Fonte: Elaborado pelos autores.

DISCUSSÃO

Os resultados revelam que a maioria dos cursos de graduação nas profissões da saúde ainda adota um paradigma tradicional de organização do processo avaliativo do estudante. Esse modelo é caracterizado pela reduzida participação institucional e pela predominância da avaliação somativa, sob responsabilidade exclusiva dos docentes ou da coordenação das disciplinas, das unidades curriculares ou dos estágios. Além disso, uma parte essencial do ciclo avaliativo – a devolutiva – parece estar desassistida, visto que o domínio das estratégias de *feedback* é limitado e o instrumental para maior investimento na avaliação formativa é pouco.

A predominância de avaliações somativas pode resultar da cultura institucional sobre avaliação, baseadas no pressuposto de que a medida final do desempenho do estudante das profissões da saúde, por meio de provas formais e suas pontuações, é indicativa de seu desempenho no porvir. Isso serviria de base para a escolha de candidatos à formação avançada, como a residência médica, e também para proteger a população dos efeitos de profissionais incompetentes. No entanto, ainda que os resultados de exames estruturados de habilidades clínicas possam identificar estudantes com risco de desempenho insuficiente no futuro, há poucas evidências de que os resultados das avaliações somativas tenham valor preditivo para esse desfecho¹⁸.

Dessa forma, a excessiva valorização da finalidade somativa costuma ocorrer em detrimento da prática da avaliação formativa, ignorando as evidências da grande efetividade da devolutiva ao estudante na promoção da aprendizagem¹⁹. Em consequência, podem estar presentes limitações no domínio de técnicas de *feedback*.

A cultura institucional no que tange à avaliação do estudante também é marcada pela escassa participação da instituição na prática da avaliação, que é delegada ao professor ou ao coordenador da unidade curricular. Isso gera enorme heterogeneidade quanto aos métodos que são utilizados na avaliação do estudante nos vários domínios, gerando dúvidas sobre a qualidade dos instrumentos e sua utilização baseada nas melhores práticas.

De fato, entre os métodos de avaliação de conhecimento declarados, predominam os testes de múltipla escolha e as provas com questões abertas (Tabela 1). Esses achados coincidem com um estudo¹⁸ que evidencia a alta utilização e a boa efetividade desses métodos para o domínio cognitivo. No entanto, dada a variedade de formatos disponíveis, é fundamental fazer escolhas criteriosas e formular as avaliações adequadamente, com base na literatura especializada, para garantir maior confiabilidade^{20,21}.

Quanto à avaliação das habilidades clínicas e procedimentais, os resultados mostraram que, além das provas

práticas e de laboratório, o exame clínico estruturado por estações (OSCE) é o mais utilizado (Tabela 2). Isso se justifica porque desenvolver essas habilidades é essencial no processo formativo dos estudantes, futuros profissionais da saúde^{22,23}. O OSCE é um modelo de avaliação que, quando bem planejado e executado, possui validade, fidedignidade e acurácia, fornecendo uma medida confiável sobre a competência em ambientes simulados^{24,25}. Ele tem um alto impacto educacional, e há evidências de associações significativamente positivas entre seus escores e o desempenho clínico, como observado em estudantes de Fisioterapia, Farmácia, Medicina, Odontologia e Nutrição¹⁸.

Embora a diversificação de métodos seja louvável, a oferta de estratégias, ora validadas, ora customizadas para determinadas atividades ou objetivos de aprendizagem, sem supervisão ou engajamento institucional, pode levar à inadequação na avaliação dos domínios. Por exemplo, alguns gestores indicaram o uso de testes de progresso ou portfólios, que são mais adequados para avaliação dos domínios cognitivo ou afetivo, respectivamente, para a avaliação de habilidades procedimentais (Tabela 2).

Avaliar o desempenho do estudante em habilidades clínicas não implica apenas detectar e corrigir deficiências, mas, principalmente, reforçar as fortalezas do desempenho do avaliado²³, o que envolve o uso de técnicas consagradas e efetivas para esse fim²⁶.

No campo do comportamento profissional, embora diversos métodos tenham sido relatados pelos gestores, alguns instrumentos são globalmente conceituados. De fato, a avaliação dos processos de aprendizagem no domínio afetivo-atitudinal costuma gerar insegurança no avaliador, pois seu julgamento pode ser mais influenciado por posições ideológicas do que na avaliação de outros domínios (cognitivo e procedimental)²⁷. Na cultura avaliativa predominante, esse quesito afetivo-atitudinal é subestimado porque não se ajusta bem à “função sancionadora quantificável” (p.208), diferentemente dos demais que se encaixam na “ilusão de se acreditar no rigor de suas afirmações porque são matematizáveis” (p. 208)²⁷.

O domínio afetivo-atitudinal contém os atributos do profissionalismo, que possui uma definição multifacetada e dinâmica, sem um conceito único na literatura²⁸. A avaliação do profissionalismo abrange diversas atitudes e comportamentos profissionais, como altruísmo, responsabilidade e trabalho em equipe, utilizando diferentes estratégias²⁸. Devido à sua importância e ao fato de que seu aprendizado não é instintivo, é fundamental incluir essa temática nos currículos das profissões da saúde e combinar métodos para uma avaliação global do estudante, além de formar os professores para isso.

Embora a utilização de diferentes métodos possa indicar heterogeneidade, cada um deles tem seus pontos fortes e limitações intrínsecas. O uso de várias estratégias de avaliação ao longo do tempo pode compensar parcialmente as deficiências de uma ou outra abordagem^{29,30}. Nesse sentido, construir sistemas de avaliação³¹ e investir na avaliação programática^{3,10,11} nos cursos pode ser uma alternativa eficaz para aprimorar os processos decisórios do progresso do estudante. Essa abordagem contínua e integrada, com registro adequado dos achados, vai além dos exames somativos, privilegiando a avaliação formativa, o *feedback* e a observação das habilidades práticas, permitindo uma compreensão abrangente e holística do percurso estudantil para adquirir as competências necessárias à futura profissão.

A mudança de paradigma no que tange à avaliação do estudante e seu contínuo aprimoramento demandam investimento permanente em DD. Embora antes se considerasse que um especialista em sua área seria naturalmente um professor eficaz, atualmente se entende que a preparação e a formação para o ensino, no exercício pleno da docência, são primordiais³².

Este artigo revela tendências de mudanças visando à melhoria dos processos avaliativos realizadas nas IES, com engajamento de professores e gestores. As atividades de DD podem acelerar transformações positivas. Considerando os professores como o recurso mais valioso da instituição, investir em seu crescimento profissional é essencial para promover inovação e excelência em todos os níveis do “*continuum* educacional”³³.

Os resultados aqui apresentados permitem identificar as características das práticas de avaliação dos estudantes em diversos cursos de graduação nas profissões da saúde, bem como as tendências de mudanças e melhorias potencializadas pelas oficinas de DD. No entanto, é importante reconhecer algumas de suas limitações.

Primeiramente, é importante considerar que, embora a proposta de DD que deu origem aos dados tenha criado uma oportunidade única para delinear diagnósticos situacionais de diversos cursos, envolvendo gestores e professores de diferentes IES, trata-se de uma amostra de conveniência. Dessa forma, apesar de oferecer um panorama relativamente amplo e diversificado, esse conjunto de dados pode não ser representativo do cenário nacional.

Além disso, o intervalo entre a coleta de dados e a divulgação pode ter ensejado outras mudanças não detectadas. Finalmente, considerando o contexto da oferta das atividades de DD, cenário para este estudo (custo baixo para as instituições participantes e custo praticamente zero para gestores e professores), relatos positivos sobre os efeitos dessas atividades podem sofrido influência desse aspecto.

Apesar dessas limitações, este estudo revela elementos de caracterização das práticas de avaliação do estudante nos cursos de graduação na área da saúde e mostra tendências de mudanças e avanços, que foram fortalecidos por atividades de DD. No entanto, cabe reconhecer que pesquisas semelhantes são desejáveis para melhor caracterizar o cenário nacional da avaliação do estudante nos cursos de graduação nas profissões da saúde e a sua transformação visando à adequação às recomendações nacionais e internacionais nesse campo.

CONSIDERAÇÕES FINAIS

Os resultados deste estudo destacam a prevalência do paradigma tradicional nos processos avaliativos dos cursos de graduação na saúde, com ênfase na função somativa em detrimento da formativa, além de revelar insuficiência na prática de *feedback* e limitada responsabilidade institucional na avaliação do estudante. Professores e gestores de diferentes instituições brasileiras mostram empenho em diversificar métodos e estratégias de avaliação, e alinhá-las às demandas profissionais, especialmente nas habilidades clínicas, um processo que parece ter sido acelerado pelas atividades de DD. No entanto, a falta de uniformidade nos métodos de avaliação do domínio afetivo-atitudinal se constitui em grande desafio. Mudanças culturais para implementar sistemas e programas de avaliação mais abrangentes são necessárias, exigindo investimento contínuo em DD, o que deve melhorar a atuação docente nas práticas educacionais nos cursos de graduação na área da saúde.

AGRADECIMENTOS

Às escolas, aos docentes, aos gestores e aos colaboradores dos cursos que participaram do “Consórcio LAG – 2018-2020 – Projeto de desenvolvimento docente para a avaliação programática”, e ao National Board of Medical Examiners (NBME).

CONTRIBUIÇÃO DOS AUTORES

Rodrigo Humberto Flauzino, Maria Paula Panúncio-Pinto, Valdes Roberto Bollela e Luiz Ernesto de Almeida Troncon participaram igualmente da concepção e do planejamento do estudo, da análise e interpretação dos dados, e da elaboração e revisão do manuscrito.

CONFLITO DE INTERESSES

Declaramos não haver conflito de interesses.

FINANCIAMENTO

Este estudo recebeu apoio financeiro da Fundação CAPES, por meio do programa PROAP/PROEX.

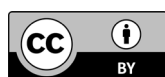
DISPONIBILIDADE DE DADOS

Os dados de pesquisa estão disponíveis no corpo do documento.

REFERÊNCIAS

1. Vianna HM. Introdução à avaliação educacional. São Paulo: Ibrasa; 1989.
2. Troncon LEA. Avaliação programática do estudante: estratégia institucional para melhor cumprir as funções da avaliação educacional. *Rev Grad USP*. 2016;1(1):53-8 [acesso em 23 jan 2024]. Disponível em: <https://www.revistas.usp.br/gradmais/article/view/117725>.
3. Troncon LEA. Estruturação de sistemas para avaliação programática do estudante de medicina. *Rev Bras Educ Med*. 2016;40(1):30-42 [acesso em 23 jan 2024]. Disponível em: <https://www.scielo.br/j/rbem/a/4g8vyJQ9Cjfr3vjgzzfhQr/?lang=pt>.
4. Palhares Neto AA, Batista NA, Diniz RV, Bollela VR. Teste de Progresso: avanços e perspectivas. *Rev Bras Educ Med*. 2022;46(supl 1):e160 [acesso em 01 dez 2023]. Disponível em: <https://www.scielo.br/j/rbem/a/rhBgjSzGbNYRjk4BcGdmLKb/?format=pdf&lang=pt>.
5. Griboski CM. Qualidade da formação dos estudantes de Medicina. *Interface (Botucatu)*. 2020;24:e190673. doi: <https://doi.org/10.1590/Interface.190673>.
6. Carmona F, Cesaretti MLR, Oliveira AS, Bollela VR. O futuro da educação na universidade: avanços possíveis e necessários. *Medicina (Ribeirão Preto)*. 2021;54(supl 1):e-189735 [acesso em 22 nov 2023]. Disponível em: <https://www.revistas.usp.br/rmrp/article/view/189735>.
7. Volpe FA, Quintana SM, Borges MC, Troncon LEA. Avaliação do estudante na educação remota (ER) e à distância (EAD): como desenvolver de modo efetivo, enfatizando a devolutiva. *Medicina (Ribeirão Preto)*. 2021;54(supl 1):e-184773 [acesso em 23 nov 2023]. Disponível em: <https://www.revistas.usp.br/rmrp/article/view/184773>.
8. International Association for Health Professions Education. The International Association for Health Education and Aspire Awards (Aspire recognition of excellence in inspirational and innovative approaches to health professions education [acesso em 31 jan 2024]. Disponível em: <https://amee.org/wp-content/uploads/2024/01/Insp-Inn-Apps-Criteria-2024.pdf>
9. World Federation for Medical Education. Basic medical education WFME global standards for quality improvement. WFME; 2020 [acesso em 23 jan 2024]. Disponível em: <https://wfme.org/wp-content/uploads/2020/12/WFME-BME-Standards-2020.pdf>.
10. Schuwirth LW, van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach*. 2011;33(6):478-85. doi: <https://doi.org/10.3109/0142159X.2011.565828>.
11. van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-14 [acesso em 23 jan 2024]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22364452/>.
12. Brasil. Resolução CNE/CNS nº 3, de 20 de junho de 2014. Institui Diretrizes Curriculares Nacionais do Curso de Graduação em Medicina e dá outras providências. *Diário Oficial da União*; 23 jun 2014. Seção 1, p. 8-11 [acesso em 23 jan 2024]. Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=15233-diretrizes-medicina&Itemid=30192.
13. Brasil. Recomendações da Comissão Intersetorial de Recursos Humanos às Diretrizes Curriculares Nacionais dos Cursos Graduação 1 – Área da Saúde. Resolução nº 569, de 8 de dezembro. 2017 [acesso em 30 nov 2023]. Disponível em: <https://conselho.saude.gov.br/resolucoes/2017/Reso569.pdf>.
14. Cintra KA, Borges MC, Panúncio-Pinto MP, Troncon LEA, Bollela VR. The impact and the challenges of implementing a faculty development program on health professions education in a Brazilian Medical School: a case study with mixed methods. *BMC Med Educ*. 2023;23:784. p. 1-9. [acesso em 23 jan 2024]. Disponível em: <https://bmcmmeduc.biomedcentral.com/articles/10.1186/s12909-023-04754-8>

15. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med.* 2006;119(2):166.e7-16 [acesso em 23 jan 2024]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/16443422/>.
16. Silva PR, Araújo ES, Caldeira AM. Construction and validation of a questionnaire for the analysis of bioethical conceptions. *Revista Bioética.* 2012;20(3):488-99 [acesso em 30 nov 2023]. Disponível em: <https://repositorium.sdum.uminho.pt/handle/1822/23304>.
17. Bardin L. Análise de conteúdo. 5a ed. Lisboa: Edições 70; 2021.
18. Terry R, Hing W, Orr R, Milne N. Do coursework summative assessments predict clinical performance? A systematic review. *BMC Med Educ.* 2017;17(1):40. p.1-15 [acesso em 30 nov 2023]. Disponível em: <https://bmcmmeduc.biomedcentral.com/articles/10.1186/s12909-017-0878-3>
19. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31, *Med Teach.* 2007;29(9):855-71 [acesso em 23 jan 2024]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/18158655/>.
20. Case SM, Swanson DB. Constructing written test questions for the basic and clinical sciences. 3rd ed. Philadelphia: National Board of Medical Examiners; 2001 [acesso em 23 jan 2024]. Disponível em: https://ia800309.us.archive.org/4/items/Constructing_Written_Test_Questions_For_the_Basic_and_Clinical_Sciences/NBMEtestItems.pdf.
21. Bollela VR, Borges MC, Troncon LEA. Avaliação somativa de habilidades cognitivas: experiência envolvendo boas práticas para a elaboração de testes de múltipla escolha e a composição de exames. *Rev Bras Educ Med.* 2018;42(4):74-85 [acesso em 23 jan 2024]. Disponível em: <https://www.scielo.br/j/rbem/a/9dnZCHRwdQKjFt7vH4DcR6n/?format=pdf&lang=pt>.
22. Harden RM. Assessment of clinical competence using objective structured clinical examination. *BMJ.* 1975;1:447-51 [acesso em 30 nov 2023]. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1672423/>.
23. Panúncio-Pinto MP, Troncon LEA. Avaliação do estudante – aspectos gerais. *Medicina (Ribeirão Preto).* 2014;47(3):314-23 [acesso em 22 nov 2023]. Disponível em: <https://www.revistas.usp.br/rmrp/article/view/86684>.
24. Davis MH. OSCE: the Dundee experience. *Med Teach.* 2003;25:255-61 [acesso em 22 nov 2023]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/12881046/>.
25. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? *Med Teach.* 2013;35(6):503-14 [acesso em 22 nov 2023]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23521582/>.
26. Bienstock JL, Katz NT, Cox SM, Hueppchen N, Erickson S, Puscheck EE. To the point: medical education reviews – providing feedback. *Am J Obstet Gynecol.* 2007;196(6):508-13 [acesso em 30 nov 2023]. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0002937806101702?via%3Dihub>.
27. Zabala A. A prática educativa: como ensinar. Porto Alegre: Artmed; 1998.
28. Carneiro M, Cunha SM, Feitosa ES, Sá RB, Brilhante AV. O profissionalismo e suas formas de avaliação em estudantes de Medicina: uma revisão integrativa. *Interface (Botucatu).* 2020;24:1-15 [acesso em 23 nov 2023]. Disponível em: <https://www.scielo.br/j/icse/a/fSVQGWk6mSKjGLyRzRXxTwH/?format=pdf&lang=en>.
29. Epstein RM. Assessment in medical education. *NEJM.* 2007;356:387-96 [acesso em 30 nov 2023]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17251535/>.
30. Al-Wardy NM. Assessment methods in undergraduate medical education. *SQUMJ.* 2010;10(2). p. 203-9. [acesso em 22 nov 2023]. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074721/>.
31. Norcini J, Anderson MB, Bollela V, Burch V, Costa MJ, Duvivier R, et al. Consensus framework for good assessment. *Med Teach.* 2018;40(11):1102-9 [acesso em 30 nov 2023]. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/0142159X.2018.1500016>.
32. Steinert Y, Mann K, Centeno A, Dolman D, Spencer J, Gelula M, et al. A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Med Teach.* 2006;28(6):497-526 [acesso em 30 nov 2023]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17074699/>.
33. Steinert Y, Mann K, Anderson B, Barnett BM, Centeno A, Naismith L, et al. A systematic review of faculty development initiatives designed to enhance teaching effectiveness: a 10-year update: BEME Guide No. 40. *Med Teach.* 2016;38(8):769-89 [acesso em 30 nov 2023]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27420193/>.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.