




Validação de conteúdo para um instrumento para avaliação de estudantes de Medicina em sessões tutoriais

Validation of content for an instrument to assess medical students in tutorial sessions

José Reinaldo Madeiro Júnior¹  jrmadeiro@uol.com.br

João Luís da Silva¹  joao.luiz@ufpe.br

Alexandre César Vieira de Sales¹  ac.sales@uol.com.br

Edvaldo da Silva Souza²  edvaldo.s@fps.edu.br

RESUMO

Introdução: O perfil do médico esperado no Brasil compreende um profissional com visão holística do ser humano. Para alcançar esse perfil, as metodologias ativas de ensino, como a aprendizagem baseada em problemas (ABP), vêm ganhando espaço. Para que a ABP atinja todos os seus objetivos na formação do profissional, uma boa avaliação é indispensável. A disponibilidade de instrumentos de avaliação validados representa um avanço na tentativa de mensurar e direcionar o aprendizado.

Objetivo: Este estudo teve como objetivos desenvolver e validar o conteúdo para um instrumento de avaliação de estudantes em sessões tutoriais para uso em cursos de Medicina que utilizem a metodologia ABP.

Método: Para a construção da versão preliminar do conteúdo destinado ao instrumento, foi realizada uma revisão sistemática rápida nas bases de dados PubMed, Ebsco e BVS. Com a revisão, desenvolveu-se a versão preliminar, que contou com 24 itens agrupados em três domínios (utilização de recursos de aprendizagem, compreensão e raciocínio – D1, profissionalismo e trabalho em equipe – D2 e resolução de problemas e efetividade no grupo – D3), cada um com oito itens. Essa versão foi encaminhada a um painel de especialistas composto por tutores com pelo menos cinco anos de experiência em ABP da Faculdade Pernambucana de Saúde (FPS), por meio da metodologia Delphi.

Resultados: Dos 32 membros inicialmente previstos no painel, 17 foram incluídos na análise dos dados, por terem respondido ao questionário completamente. Na primeira rodada, alcançou-se um valor de concordância parcial ou total superior a 70%, valor inicialmente previsto para todos os 24 itens do instrumento, o que dispensou a necessidade de uma segunda rodada. Para diminuir o número final de itens, optou-se por manter apenas os itens com pelo menos 70% de concordância total, tendo a versão final do instrumento quatro itens no D1, cinco no D2 e quatro no D3.

Conclusões: O instrumento foi validado com um total de 13 itens. O conteúdo para o instrumento apresentou componentes com várias semelhanças em relação aos encontrados nos instrumentos publicados na literatura e já validados, que, por sua vez, estão de acordo com os objetivos de aprendizagem propostos pela ABP.

Palavras-chave: Avaliação Educacional; Tutoria; Educação Médica; Aprendizagem Baseada em Problemas.

ABSTRACT

Introduction: The expected doctor profile in Brazil comprises a professional with a holistic view of the human being. To achieve this profile, active teaching methodologies, such as Problem-Based Learning (PBL), are gaining ground. For PBL to achieve all its objectives in the training of professionals, good assessment is essential. The availability of validated assessment instruments represents an advance in the attempt to measure and direct learning.

Objective: The aim of this study was to develop and validate the content for an instrument for assessing students in tutorial sessions for use in medical courses that use PBL methodology.

Method: For the construction of the preliminary version of the content for the instrument, a Rapid Systematic Review was carried out in the PUBMED, EBSCO and BVS databases. The preliminary version was thus developed with 24 items grouped in three domains (use of learning, understanding and reasoning resources - D1, Professionalism and teamwork - D2, and Problem-solving and group effectiveness - D3), each with eight items. This version was sent to a panel of experts made up of tutors with at least five years' experience in PBL at the Faculdade Pernambucana de Saúde (FPS), using the Delphi methodology.

Results: Of the 32 members initially earmarked for the panel, 17 were included in the data analysis, for having fully answered the questionnaire. In the first round, a partial or total agreement value greater than 70 percent was reached, which was initially expected, for all 24 items of the instrument, which dispensed with the need for a second round. To reduce the final number of items, only those with at least 70 percent total agreement were maintained, with the final version of the instrument containing four items on D1, five on D2 and four on D3.

Conclusions: The instrument was validated with a total of thirteen items. The content for the instrument brought components with several similarities to those that make up the instruments published in the literature that have already been validated, which in turn are in accordance with the learning objectives proposed by PBL.

Keywords: Educational Assessment; Mentoring; Medical Education; Problem-Based Learning.

¹ Universidade Federal de Pernambuco, Caruaru, Pernambuco, Brasil.

² Faculdade Pernambucana de Saúde, Recife, Pernambuco, Brasil.

Editora-chefe: Daniela Chiesa | Editora associada: Daniela Chiesa

Recebido em 06/08/20; Aceito em 26/04/21. | Avaliado pelo processo de *double blind review*.

INTRODUÇÃO

A partir da década de 1970, aprofundou-se no Brasil e em toda a América Latina a discussão sobre como tornar a formação médica congruente com a característica do profissional necessário para atender às demandas da sociedade. O modelo flexneriano, caracterizado por fragmentação do ensino e tendência à hiperespecialização, num contexto em que a formação se dá basicamente nos hospitais escola, não vinha conseguindo atender a essa demanda, apesar de ainda hoje estar bastante presente¹.

Seguindo esse raciocínio, tem sido descrito que as metodologias ativas de ensino aprendizagem apresentam uma oportunidade de modificar a formação médica no sentido de gerar o profissional necessário. Como características dessas metodologias, temos a maior integração com a realidade sociocultural na qual o estudante está inserido e a necessidade de construir o conhecimento coletivamente, valorizando conhecimentos prévios. Além disso, oferecem a possibilidade de formar um profissional capaz de ter as ferramentas necessárias para o constante aprimoramento profissional, tão necessário em uma realidade em que o conhecimento médico cresce exponencialmente².

A aprendizagem baseada em problemas (ABP) é uma das mais difundidas metodologias ativas de ensino. Na ABP, um dos ambientes de aprendizagem mais importantes é a sessão tutorial, que se caracteriza pela presença de grupos em média de oito a dez estudantes que seguem passos preestabelecidos para a construção do conhecimento. Na ABP, o tutor tem um papel que é bastante diferente do papel docente nas metodologias tradicionais de ensino, em que predomina a transmissão de conhecimentos. Um bom tutor deve ser capaz de motivar seus estudantes a trabalhar em equipe no intuito de solucionar situações-problema colocadas a eles, direcionando-os a encontrar o melhor caminho para gerar a produção do conhecimento, num papel de suporte conhecido como *scaffolding*³⁻⁴.

O tutor tem um papel essencial para o bom funcionamento de um grupo tutorial. Entretanto, como boa parte dos tutores foi formada na metodologia tradicional, isso pode causar-lhes um certo desconforto ao exercerem a função e levá-los a não exercer o papel de dar o verdadeiro suporte para o aluno construir o próprio conhecimento. Dessa maneira, encontramos tutores com papel excessivamente passivo, que permanecem praticamente alheios ao processo. Há outros que, ao sentirem uma preocupação excessiva na cobertura completa dos conteúdos propostos, monopolizam o processo e dificultam que o aluno exerça sua desejada autonomia⁵.

Como condição indispensável para que o tutor exerça seu papel adequadamente, está a avaliação que é realizada nesses estudantes. O reconhecimento de que a avaliação

direciona o processo de aprendizagem e não deve ter caráter meramente somativo, mas ser um processo contínuo de ajuste de rumo, faz com que se busquem novas formas de avaliar. Essa reflexão, porém, torna ainda mais complexa e difícil a prática avaliativa¹.

Quando se realiza uma avaliação, vem à tona a necessidade de discutir o conceito de competências, indo além da simples noção de saber fazer uma determinada atividade. É necessário inferir a capacidade do sujeito de discernir as diferentes situações em que estão inseridas essas ações. Dessa forma, ultrapassaremos o “saber fazer” e chegaremos ao “saber agir”⁶.

Para que a avaliação de discentes em sessões tutoriais dentro da metodologia ABP alcance seus objetivos, é necessário atentar ao fato que existem características como a subjetividade humana e a limitação do tempo, que nem sempre estão presentes em outras formas de avaliação⁷. Vários métodos diferentes já foram utilizados, desde os mais tradicionais, como testes de múltipla escolha, até formas mais recentes, como autoavaliação e avaliação entre pares. Uma metanálise realizada em 2005 ressaltou a importância da congruência entre as atividades desenvolvidas na ABP e a avaliação⁸.

Para alcançar esse objetivo, o uso de instrumentos de avaliação é extremamente relevante. Para que possa atingir esse propósito, um instrumento deve ser válido, o que significa ser capaz de efetivamente mensurar aquilo a que se propõe. Precisa também apresentar confiabilidade, que está relacionada à reprodutibilidade dos escores que são obtidos de uma avaliação. Além desses dois tradicionais critérios, podemos incluir viabilidade, efeito educacional e aceitabilidade como importantes características de um instrumento. A viabilidade refere-se a custo acessível. O efeito educacional compreende o direcionamento da motivação dos estudantes para o cumprimento de um determinado objetivo, e a aceitabilidade significa que o público interessado endossa a proposta de avaliação⁹.

Os instrumentos de avaliação já reportados na literatura compartilham a intenção de avaliar a *performance* discente em domínios como capacidade de raciocínio, aplicação de conhecimento, resolução de problemas, profissionalismo, trabalho em grupo e estudo independente. Esses estudos foram realizados na escola de origem dos autores, não havendo publicação de trabalhos que englobem diferentes países ou regiões¹⁰⁻¹⁵.

Observa-se que os instrumentos validados utilizados para avaliação de estudantes em sessões tutoriais não apresentam divergências, principalmente no tocante à quantidade de itens, sendo necessária a realização de novos estudos para estabelecer que itens devem compor um instrumento que irá avaliar da melhor maneira possível a

performance dos discentes. Salienta-se também a importância de que tais instrumentos sejam elaborados de acordo com a realidade de cada local, respeitando-se, assim, particularidades culturais, socioeconômicas e epidemiológicas, entre outras.

Verifica-se a escassez de estudos em língua portuguesa, com incorporação de características culturais da região, o que justifica o desenvolvimento de um novo instrumento.

Este estudo teve como objetivos desenvolver e realizar a validação de conteúdo para um instrumento de avaliação de estudantes de Medicina em sessões tutoriais, em língua portuguesa.

MÉTODO

O estudo foi realizado em duas etapas. Inicialmente, fez-se uma revisão sistemática rápida, baseada no protocolo produzido pela plataforma PRISMA, com o objetivo de identificar as competências que devem ser avaliadas em estudantes de Medicina, em sessões tutoriais. Elegeram-se artigos, em inglês ou português, que apresentavam instrumentos de avaliação de estudantes de Medicina em sessões tutoriais ou que descrevessem competências que estes devem ter em sessões tutoriais na metodologia ABP.

A pesquisa foi realizada nas bases de dados PubMed, Ebsco e BVS, no período de 2 a 28 de dezembro de 2017 e de 4 a 31 de janeiro de 2018, sem restrição de ano de publicação. Além disso, fez-se também uma pesquisa manual nas referências bibliográficas dos artigos selecionados.

Para a base de dados PubMed, empregou-se a seguinte combinação de palavras-chave (*Mesh terms*): *tutorial AND assessment OR evaluation OR competencies AND medical education AND problem based learning*, sendo considerados apenas artigos em inglês e português, no período de 2 a 18 de dezembro de 2017.

Para a base de dados Ebsco, a combinação de palavras-chave foi: *tutor OR tutorial AND assessment OR evaluation AND problem based learning AND medical education OR medical school OR medical students OR medical curriculum*, no período de 4 a 31 de janeiro de 2018.

Já na base de dados BVS, utilizaram-se as seguintes palavras-chave *avaliação AND tutoria AND aprendizado baseado em problemas OR PBL*, no período de 28 a 31 de janeiro de 2018.

Para a validação do conteúdo para o instrumento identificado na revisão sistemática rápida, utilizou-se o método Delphi, que se caracteriza por realizar um consenso entre *experts* no tema¹⁶. Para a composição do painel de especialistas, optou-se por selecionar docentes da Faculdade Pernambucana de Saúde (FPS) com pelo menos cinco anos de atuação como tutores na metodologia ABP. Dessa forma, garantiu-se que o corpo de especialistas tivesse maturidade e experiência

profissionais suficientes. Foram excluídos tutores participantes da pesquisa e tutores afastados para qualificação profissional, por licença-maternidade e por problemas de saúde. A FPS possui um curso de Medicina que oferece 192 vagas por ano, e oito turmas já o concluíram. A metodologia utilizada é, desde o início do curso, a ABP. A instituição também oferece mestrado profissional em Educação para o Ensino na Área de Saúde e em Psicologia da Saúde. O programa começou em 2011 e utiliza metodologias ativas de ensino.

O conteúdo para o instrumento, em sua versão preliminar, foi enviado por correspondência eletrônica ao painel de especialistas. Os itens foram apresentados em formato de escala de Likert, sendo solicitado o grau de concordância com a presença do item na versão final do instrumento: 1 = concordo totalmente, 2 = concordo parcialmente, 3 = não concordo nem discordo, 4 = discordo parcialmente e 5 = discordo totalmente.

Para o envio do referido conteúdo, utilizou-se o *software* LimeSurvey. Neste estudo, adotou-se o valor de 70% de consenso na concordância total ou parcial de um determinado item como nível mínimo a ser alcançado para a inclusão do item na versão final do conteúdo para o instrumento, valor comumente utilizado em estudos similares. Previu-se um total de duas ou três rodadas.

Para a avaliação dos estudantes, é utilizado atualmente na FPS um instrumento que inclui os seguintes critérios: pontualidade; uso do conhecimento prévio para explicar e identificar o problema; participação no fórum do ambiente virtual de aprendizagem entre as sessões de abertura e fechamento, e contribuição com informações para o estudo do grupo; exposição de ideias de forma clara e organizada no grupo; bom desempenho da função de coordenador, secretário ou membro do grupo; interação harmônica com os demais membros do grupo. A avaliação é realizada ao final de cada grupo tutorial.

A coleta de dados iniciou-se em outubro de 2018, quando da primeira rodada Delphi, após aprovação do Comitê de Ética em Pesquisa da FPS, e estendeu-se até janeiro de 2019.

RESULTADOS

Na revisão sistemática rápida, consultaram-se três bases de dados. Na base de dados PubMed, identificaram-se inicialmente 1.637 artigos, e mais um foi identificado na verificação manual das referências dos artigos selecionados. Desses, 1.571 foram descartados pelo título, por não fazerem referência ao tema da pesquisa. Dos 66 artigos restantes, descartaram 42 após leitura do resumo e 17 depois da leitura na íntegra. Por fim, incluíram-se sete artigos.

Quanto à base de dados Ebsco, inicialmente se identificaram 3.865 artigos, sendo 3.841 descartados pelo

título, por não se referirem ao tema da pesquisa. Dos 24 restantes, dez foram descartados por já terem sido identificados na pesquisa feita na plataforma PubMed e 12 após leitura do resumo. Finalmente, dois artigos foram lidos na íntegra, não sendo nenhum deles incluído no estudo. Por fim, na base de dados BVS, foram inicialmente identificados 199 trabalhos, sendo 178 descartados pelo título e 21 pelo resumo, todos por não fazerem referência ao objeto desta pesquisa.

Ao final da revisão, identificaram-se sete artigos científicos que traziam instrumentos de avaliação (Quadro 1)^{10-15, 17}. Após, elaborou-se uma versão inicial do conteúdo para o

instrumento, baseado na síntese dos instrumentos analisados. Buscou-se incorporar pontos de todos os instrumentos selecionados, criando domínios que resumissem os principais tópicos presentes nos instrumentos identificados. Para não tornar o número de itens extremamente elevado, retiraram-se aqueles que estariam contemplados em outros, sem prejuízo do conteúdo final. Finalmente, a versão inicial do conteúdo para o instrumento foi composto por itens divididos nos domínios de aprendizagem utilização de recursos de aprendizagem, compreensão e raciocínio, profissionalismo e trabalho em equipe, e resolução de problemas e efetividade no grupo, cada um com oito itens (Quadro 2).

Quadro 1. Estudos sobre instrumentos de avaliação de estudantes em sessões tutoriais na metodologia ABP – Recife, Pernambuco, Brasil, 2018.

AUTOR, ANO E PAÍS DE ORIGEM DO ESTUDO	CONTEÚDO PRÉVIO DO INSTRUMENTO	FORMA DE APLICAÇÃO	FORMATO DA ANÁLISE DOS DADOS	RESULTADOS PRINCIPAIS	LIMITAÇÕES E SUGESTÕES DESCRITAS
Hébert et al. (1996), Canadá	Composto por 44 itens, o Tutotest identifica a frequência de um determinado comportamento. Os itens foram elaborados por professores e estudantes da Universidade de Sherbrooke.	Aplicaram-se 270 avaliações em 100 estudantes ao fim das quatro unidades do primeiro ano.	<ul style="list-style-type: none"> Análise fatorial exploratória. Determinação do coeficiente alfa de Cronbach. Determinação do coeficiente de correlação intraclasse (CCI). Correlação com resultados de testes escritos. 	<ul style="list-style-type: none"> Itens agrupados em quatro domínios: efetividade no grupo, habilidades de comunicação e liderança, curiosidade científica e respeito aos colegas. Boa correlação com avaliação global do tutor. Moderada correlação com testes escritos; Confiabilidade determinada pelo alfa de Cronbach: 0,98. CCI de 0,81 após cinco aplicações do Tutotest. Moderada correlação com resultados de testes escritos. 	Necessidade de novos estudos para verificar a estrutura fatorial e identificar itens que poderiam ser removidos.
Valle et al. (1999), México	Docentes, psicólogos e especialistas em educação da Universidade Autônoma do México elaboraram 39 itens, distribuídos nos domínios: estudo independente, interação no grupo e habilidades de raciocínio. Considerados não relevantes, dez itens foram excluídos.	Em 16 grupos tutoriais, 14 tutores avaliaram 152 estudantes do primeiro e segundo anos.	<ul style="list-style-type: none"> Análise fatorial exploratória. Determinação do coeficiente alfa de Cronbach. 	<ul style="list-style-type: none"> Distribuíram-se 24 itens em quatro domínios: estudo independente, interação em grupo, habilidades de raciocínio e participação ativa. Excluíram-se cinco itens excluídos por baixa correlação após análise fatorial exploratória. Confiabilidade verificada no alfa de Cronbach: 0,96. 	Indicação de novos estudos para realizar a correlação entre performance na tutoria e habilidades profissionais requeridas.
Elizondo-Montemayor (2004), México	Agruparam-se 50 itens em “rubricas” derivadas dos objetivos principais da ABP (aplicação do conhecimento, raciocínio crítico, estudo autodirecionado e colaboração), acrescidos de comportamento profissional, em escala de seis itens.	Instrumento utilizado nas disciplinas do curso básico e na disciplina de ginecologia e obstetrícia da Escola Médica Tec. de Monterrey.	Não foi realizada.	<ul style="list-style-type: none"> De acordo com o feedback dos tutores, o instrumento colaborou para a demonstração da performance dos estudantes. 	Não descritas

Continua...

Quadro 1. (Continuação) Estudos sobre instrumentos de avaliação de estudantes em sessões tutoriais na metodologia ABP – Recife, Pernambuco, Brasil, 2018.

AUTOR, ANO E PAÍS DE ORIGEM DO ESTUDO	CONTEÚDO PRÉVIO DO INSTRUMENTO	FORMA DE APLICAÇÃO	FORMATO DA ANÁLISE DOS DADOS	RESULTADOS PRINCIPAIS	LIMITAÇÕES E SUGESTÕES DESCRITAS
Chen (2006), Taiwan	Após revisão bibliográfica e entrevistas com especialistas no tema, elaboraram-se 15 itens, que foram divididos em quatro categorias: senso crítico, utilização de fontes de aprendizagem, trabalho em equipe e atitudes e habilidades em comunicação, em escala de Likert de cinco itens.	Participaram 136 estudantes do pré-internato, divididos em 18 itens e avaliados por 47 tutores.	• Determinação do coeficiente alfa de Cronbach.	• Escore de 0,97 do alfa de Cronbach, demonstrando boa confiabilidade do instrumento	Necessidade de utilização do instrumento por anos seguidos, para obter melhor análise e permitir o aprimoramento dele.
Sim et al. (2006), Malásia	Avaliação de quatro habilidades (participação e habilidades de comunicação, cooperação e trabalho em equipe, compreensão e habilidades de raciocínio, conhecimentos e habilidades de reter informações), em escala de cinco itens, variando de insatisfatório a excelente, desenvolvidos por três dos autores mais um visitante.	Aplicado durante dois anos a alunos que cursavam os dois primeiros anos do curso médico da Universidade da Malásia, avaliados por 82 tutores.	• Determinação do coeficiente de confiabilidade a partir do CCI, estabelecido por meio das avaliações de diferentes tutores a respeito do aluno.	• O coeficiente de confiabilidade demonstrou confiabilidade. • O questionário específico respondido por 34 tutores demonstrou boa aceitação do instrumento.	Mais benefícios serão alcançados após mais treinamento no uso do instrumento e entendimento sobre seus propósitos por parte dos tutores.
Leung (2008), Taiwan	Versão chinesa do Tutotest de 44 itens, divididos em cinco categorias: efetividade no grupo, domínio do método, comunicação e liderança, curiosidade científica e respeito aos colegas.	Aplicado em alunos do segundo, terceiro e quarto anos do curso de Medicina da Taiwan University Hospital, totalizando 370 avaliações realizadas por 44 tutores.	• Análise fatorial exploratória. • Verificação de confiabilidade por aplicação do alfa de Cronbach e do teste-reteste duas semanas após.	• Itens agrupados em quatro domínios: efetividade no grupo, comunicação e liderança, respeito aos colegas e formação e testagem de hipóteses. • Alfa de Cronbach: 0,97. • Correlação do teste-reteste em duas semanas: 0,85.	• Um grupo de estudantes avaliados por um único tutor ao longo do período letivo pode gerar bias. • Necessidade de estudos longitudinais no intuito de verificar mudanças na performance dos estudantes ao longo do tempo.
Lee et al. (2016), Estados Unidos	Instrumento baseado em quatro competências (resolução de problemas, uso de informação, trabalho em equipe e profissionalismo) e elaborado por um painel de especialistas	Aplicado em 310 estudantes dos dois primeiros anos da Universidade da Califórnia, divididos em duas coortes, sendo cada aluno avaliado por apenas um tutor.	• Adequação aos padrões de testes das Standards for educational and psychological testing para estabelecimento da validade de conteúdo. • Determinação da confiabilidade pelos coeficientes de generalizabilidade e dependabilidade.	• Correlação positivas com seis outros testes já validados e amplamente utilizados na educação médica dos Estados Unidos.	• Estudo realizado em uma instituição única, com desenho metodológico particular, o que pode limitar sua extensão a outras escolas. • Incapacidade do estudo de concluir sobre a validade de construto.

Fonte: Elaborado pelos autores.

Quadro 2. Versão preliminar do conteúdo para o instrumento – Recife, Pernambuco, Brasil, 2018.

Domínio 1 UTILIZAÇÃO DE RECURSOS DE APRENDIZAGEM, COMPREENSÃO E RACIOCÍNIO	Domínio 2 PROFISSIONALISMO-TRABALHO EM EQUIPE	Domínio 3 RESOLUÇÃO DE PROBLEMAS- EFETIVIDADE NO GRUPO
É capaz de formular conclusões sobre o problema.	Comunica aos colegas quando se sente desconfortável com o andamento da discussão.	Esclarece, define e analisa a situação-problema no caso.
Usa os dados essenciais fornecidos pelo caso.	Identifica e expõe respeitosamente discordâncias diante das opiniões dos seus pares.	Propõe hipóteses diagnósticas e tratamentos.
Identifica falhas fornecidas nos dados oferecidos pelo caso.	Solicita feedback ao grupo.	Utiliza e oferece ao grupo referências bibliográficas adequadas para a resolução da situação-problema.
Demonstra conhecimento prévio referente à situação-problema.	Aceita sugestões.	Propõe questionamentos que ajudam a esclarecer pontos obscuros.
Participa ativamente na definição dos objetivos de aprendizagem.	Aceita decisões tomadas pelo grupo.	Lidera o grupo na discussão.
Demonstra esforços para melhorar.	É atento à participação dos outros membros do grupo.	Preocupa-se com as limitações de tempo.
Procura aconselhamento dos colegas.	Estimula a participação dos colegas.	Demonstra discordância quando um membro ou o tutor monopolizam a discussão.
Identifica situações de oportunidades para aquisição de novos conhecimentos.	Respeita o direito de outros colegas se pronunciarem.	Ajusta-se a diferentes papéis no grupo.

Fonte: Elaborado pelos autores.

O conteúdo para o instrumento, em sua versão preliminar, foi enviado por correspondência eletrônica aos 32 tutores da FPS com pelo menos cinco anos de atuação. Do total de tutores, 20 aceitaram participar da pesquisa, dos quais três foram excluídos por não terem preenchido o questionário completo. Dos 17 participantes, cinco eram do sexo masculino, e 12, do sexo feminino. O tempo médio de formado em medicina foi de 29 anos, variando de 14 a 36. O tempo de exercício na função docente em uso da metodologia ABP variou de cinco a 12 anos. Entre os participantes, 11 tinham doutorado e seis eram mestres, sendo dois com mestrado na área de educação. Dos 17, sete eram membros do Núcleo Docente Estruturante do curso, e dois eram docentes também do mestrado em Educação para o Ensino na Área de Saúde da instituição.

Verificou-se nível de concordância total ou parcial superior a 70% para todos os itens da versão preliminar do conteúdo para o instrumento. Assim, concluída a análise dos dados dessa primeira etapa, optou-se por incluir no instrumento apenas os itens que apresentaram 70% ou mais de concordância total sobre cada item. Retiraram-se quatro itens do domínio utilização de recursos de aprendizagem, três do domínio profissionalismo e trabalho em equipe e quatro do domínio resolução de problemas e efetividade no grupo. Diante do alto grau de concordância verificado, foi considerada

desnecessária a realização de uma segunda rodada de apreciação do instrumento pelos especialistas (Tabela 1).

DISCUSSÃO

A avaliação como direcionadora do processo de aprendizagem tem sido considerada um importante ponto de discussão na educação médica, buscando ultrapassar a concepção de avaliação *da* aprendizagem para alcançar uma avaliação *para* a aprendizagem¹⁸. Nesse contexto, a elaboração de instrumentos validados oferece a possibilidade tornar a avaliação mais fidedigna e capaz de alcançar seus objetivos.

Os instrumentos selecionados na revisão sistemática trazem itens que não necessariamente são de fácil interpretação, o que caracteriza uma dificuldade para o grupo de *experts*. Como exemplo, pode-se analisar o item “Demonstra esforços para melhorar”, que permite o entendimento sobre a referência a uma melhora puramente individual ou uma evolução na capacidade coletiva do grupo tutorial em alcançar os objetivos de aprendizagem.

O conteúdo para o instrumento validado neste estudo, com um alto grau de concordância entre os painelistas já na primeira rodada, foi realizado em um contexto sociocultural específico, diferente dos locais sedes dos outros estudos já publicados. Entretanto, isso não representou diferenças substanciais dos instrumentos já publicados na literatura^{10,12-15}.

Tabela 1. Versão final do conteúdo para o instrumento com o percentual de concordância total entre os *experts* respondedores – Recife, Pernambuco, Brasil, 2018.

Domínios	Concordância total entre os experts (%)
Domínio 1 – Utilização de recursos de aprendizagem, compreensão e raciocínio	
É capaz de formular conclusões sobre o problema.	82,3 %
Usa os dados essenciais fornecidos pelo caso.	82,3 %
Participa ativamente na definição dos objetivos de aprendizagem.	76,5 %
Demonstra esforços para melhorar.	82,3 %
Domínio 2 – Profissionalismo-trabalho em equipe	
Identifica e expõe respeitosamente discordâncias diante das opiniões dos seus pares.	76,5 %
Aceita sugestões.	82,3%
É atento à participação dos outros membros do grupo.	70,6 %
Estimula a participação dos colegas.	70,6%
Respeita o direito de outros colegas se pronunciarem.	88,2 %
Domínio 3 – Resolução de problemas-efetividade no grupo	
Esclarece, define e analisa a situação-problema no caso.	82,3 %
Propõe questionamentos que ajudam a esclarecer pontos obscuros.	76,5%
Preocupa-se com as limitações de tempo.	70,6%
Ajusta-se a diferentes papéis no grupo.	70,6%

Fonte: Elaborada pelos autores.

Atribui-se isso ao fato de a versão preliminar do conteúdo para o instrumento apresentada aos especialistas ter contemplado itens já descritos na literatura, que estão de acordo com os instrumentos anteriores e com os objetivos de aprendizagem da ABP, e ao fato de o painel ter sido formado por tutores experientes, com no mínimo cinco anos de atuação na área.

O número de itens da versão final desse instrumento foi de 13, tendo sido verificados na literatura instrumentos que variaram de quatro a 50 itens. Por um lado, um instrumento muito extenso pode tornar a avaliação cansativa, exigir muito tempo para sua aplicação e comprometer sua viabilidade. Por outro lado, uma quantidade maior de itens pode representar uma capacidade de abranger maior número de competências que poderiam ser esperadas do estudante.

No intuito de contribuir para a discussão quanto ao número de itens, podemos fazer uma análise mais pormenorizada de itens excluídos da versão final do conteúdo para esse instrumento. No domínio 1, notemos a exclusão do item “Demonstra conhecimento prévio referente à situação-problema”. É inerente ao arcabouço epistemológico da ABP a percepção de que o aluno deve construir seu conhecimento a partir de situações-problema tendo como ponto de partida conhecimentos prévios, sejam individuais ou adquiridos dos demais membros do grupo. Dessa forma, a exclusão desse item não pode significar uma negativa a esse paradigma. O que se espera é a interpretação por parte do avaliador que essa característica pode contribuir para a melhor *performance* em outros itens, como os itens “É capaz de

formular conclusões sobre o problema”, “Usa os dados essenciais fornecidos pelo caso” e “Participa ativamente na definição dos objetivos de aprendizagem”.

Já no domínio 3, de maneira similar, a exclusão do item “Utiliza e oferece ao grupo referências bibliográficas adequadas para a resolução da situação-problema” não significa que a capacidade de um estudante em buscar conhecimento nas melhores fontes de aprendizagem não deva ser valorizada como algo crucial para o desempenho dele, particularmente no cenário médico-científico atual de enorme divulgação de informações, algumas delas de origem não confiável. Espera-se, portanto, que itens como “Propõe hipóteses diagnósticas e tratamentos” serão bem avaliados quando o estudante os propuser respaldado nas melhores evidências científicas.

O instrumento atualmente utilizado na instituição-sede deste estudo apresenta apenas seis itens. Entre eles, podemos citar o item pontualidade, que não esteve presente em nenhum instrumento selecionado na revisão sistemática rápida, sendo apenas citado como relevante para a avaliação do critério profissionalismo no estudo de Lee et al.¹⁵. Outro item a destacar é a participação no fórum do ambiente virtual de aprendizagem. Os estudos selecionados não descrevem a presença ou não desse ambiente de aprendizagem nas suas escolas.

Além de não contemplar esses dois pontos anteriormente citados, o conteúdo validado neste estudo mostra uma presença maior de itens que serão mais observáveis na primeira sessão tutorial, em detrimento de

itens que avaliem questões mais relacionadas à sessão de fechamento. Assim, características como as relacionadas a aspectos cognitivos, à capacidade de síntese e à aquisição de conhecimentos podem não ser avaliadas satisfatoriamente.

Verifica-se, na maioria das escolas médicas brasileiras que utilizam a ABP, a presença de docentes com excelente formação técnica na área de saúde, mas nem sempre com a formação pedagógica adequada. Essa deficiência na formação docente pode fazer com que alguns entendam a avaliação apenas como função burocrática para definição de critérios com o propósito de selecionar os alunos que irão avançar no curso, não considerando toda a magnitude da avaliação como direcionadora da aprendizagem. Dificilmente um docente fará uma boa avaliação sem total conhecimento dos seus propósitos. Tem sido descrito que alguns tutores, ao não compreenderem a complexidade da ABP, privilegiam na avaliação apenas aspectos cognitivos, em detrimento de aspectos como trabalho em equipe, cooperação e construção coletiva do conhecimento¹⁹.

Outro aspecto que deve ser pontuado é que a grande maioria dos docentes teve sua formação na metodologia tradicional baseada em aulas expositivas e memorização de conteúdos. Esse aspecto faz com que alguns não exerçam adequadamente sua função de tutor, tendo uma excessiva preocupação com a cobertura dos conteúdos. Isso coíbe o protagonismo do estudante e o impede de exercer seu potencial em questões como liderança, trabalho em equipe e estímulo à participação dos colegas. Sem espaço para desenvolverem essas habilidades e atitudes, os alunos não poderão ser avaliados adequadamente em todos os aspectos do instrumento⁴.

É muito importante que seja realizada, entre os tutores da instituição que utilizam um instrumento de avaliação, a uniformização sobre como interpretar cada item, o que, por vezes, pode ser bastante complexo. Podemos citar, por exemplo, a dificuldade de conceituar uniformemente o termo profissionalismo, que, por vezes, mostrou-se descrito de formas diferentes na literatura²⁰. No conteúdo para instrumento validado neste estudo, observamos itens como “Identifica situações de oportunidades para aquisição de novos conhecimentos”, que pode não ser facilmente observável ou ser interpretado de forma desigual por parte dos tutores, ou “Usa os dados essenciais fornecidos pelo caso”, que pode levar à conclusão, às vezes equivocada, de que os tutores têm a mesma clareza sobre quais são esses dados, a partir da noção de que estes serão os tópicos do caso mais importantes para a resolução da situação-problema.

CONCLUSÃO

Observou-se que a revisão sistemática previamente realizada identificou um número pequeno de instrumentos

de avaliação. Esse fenômeno pode ter ocorrido porque muitas instituições de ensino usuárias da metodologia ABP não possuem seus instrumentos de avaliação publicados na literatura. É de se esperar que, se mais instrumentos estivessem disponíveis, novas informações e novos itens poderiam ter sido avaliados pelos *experts* como importantes para a composição dessa versão final. Além disso, não fica claro se a exclusão de determinados itens da versão inicial foi baseada na necessidade de diminuir o tamanho do instrumento ou na percepção de alguns *experts* que tais competências seriam dispensáveis.

Conclui-se que o conteúdo para o instrumento apresentou validade de conteúdo, com alto percentual de concordância, de um instrumento que engloba os domínios utilização de recursos de aprendizagem, compreensão e raciocínio (quatro itens), profissionalismo e trabalho em equipe (cinco itens) e resolução de problemas e efetividade no grupo (quatro itens), o que está congruente com os objetivos da metodologia ABP descritos na literatura

Este estudo tem como outra limitação a característica de ter validado um instrumento com a participação de tutores de apenas uma instituição, que tem missão e cultura organizacional diferentes de outras. Além disso, necessitamos da utilização do instrumento por um período considerável de tempo para atestar sua capacidade de direcionar a aprendizagem dos estudantes. Sugerem-se também novos estudos que avaliem a possibilidade de refinar ainda mais o instrumento, diminuindo a quantidade de itens, aumentando assim sua viabilidade e aceitabilidade por parte dos avaliadores, sem com isso diminuir a importância de considerar a avaliação de alunos em tutorias em toda a sua magnitude.

CONTRIBUIÇÃO DOS AUTORES

José Reinaldo Madeiro Júnior participou da construção do estudo no que se refere à concepção do estudo, da coleta e análise dos dados, da redação do manuscrito e da revisão da versão final. João Luís da Silva e Edvaldo da Silva Souza contribuíram com a revisão final. Alexandre César Vieira de Sales participou da concepção do estudo e da análise dos dados.

CONFLITO DE INTERESSES

Declaramos não haver conflito de interesses.

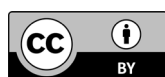
FINANCIAMENTO

Declaramos não haver financiamento.

REFERÊNCIAS

1. Bracialli LAD, Oliveira MAC. Desafios na formação médica: a contribuição da avaliação. *Rev Bras Educ Med*. 2012;36(2):280-8. doi: 10.1590/S0100-55022012000400018.

2. Mitre SM, Siqueira-Batista R, Girardi-de-Mendonça JM, Morais-Pinto NM, Meirelles CAB, Pinto-Porto C, et al. Metodologias ativas de ensino-aprendizagem na formação profissional em saúde: debates atuais. *Ciênc Saúde Colet*. 2008;13(2):2133-44. doi: 10.1590/S1413-81232008000900018.
3. Fernandes JD, Ferreira SL, La Torre MPS, Santa Rosa DO, Costa HOG. Estratégias para a implantação de uma nova proposta pedagógica na escola de enfermagem da Universidade Federal da Bahia. *Rev Bras Enferm*. 2003;56(4):392-5. doi: 10.1590/S0034-71672003000400017.
4. Martins AC, Falbo Neto G, Silva FAM. Características do tutor efetivo em ABP: uma revisão de literatura. *Rev Bras Educ Med*. 2018;42(1):104-11. doi: 10.1590/1981-52712018v42n1rb20160100.
5. Dolmans DHJM, Gijsselaers WH, Moust JHC, de Grave WS, Wolfhagen IHAP, Van der Vleuten CPM. Trends in research on the tutor in problem-based learning: conclusions and implications for educational practice and research. *Med Teach*. 2002;24(2):173-80. doi:10.1080/01421590220125277.
6. Gontijo ED, Alvim C, Megale L, Melo JRC, Lima MECC. Matriz de competências essenciais para a formação e avaliação de desempenho de estudantes de medicina. *Rev Bras Educ Med*. 2013; 37(4):526-39. doi: 10.1590/S0100-55022013000400008.
7. Waas V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet*. 2001; 357(9260):945-9. doi: 10.1016/S0140-6736(00)04221-5.
8. Gijbels D, Dochy F, Van der Bossche P, Segers M. Effects of problem-based learning: a meta-analysis from the angle of assessment. *Rev Educ Res*. 2005;75(1):27-61. doi: 10.3102/00346543075001027.
9. Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-17. doi: 10.1111/j.1365-2929.2005.02094.x.
10. Hébert R, Bravo G. Development and validation of an evaluation instrument for medical students in tutorials. *Acad Med*. 1996;71(5):488-94. doi: 10.1097/00001888-199605000-00020.
11. Leung KK, Wang WD. Validation of the Tutotest in a hybrid problem-based learning curriculum. *Adv Health Sci Educ Theory Pract*. 2008;13(4):469-77. doi: 10.1007/s10459-007-9059-1.
12. Valle R, Petra L, Martinez-Gonzalez A, Rojas-Ramirez JA, Morales-Lopez S, Piña-Garza B. Assessment of student performance in problem-based learning tutorial sessions. *Med Educ*. 1999; 33(11):818-22. doi: 10.1046/j.1365-2923.1999.00526.x.
13. Elizondo-Montemayor LL. Formative and summative assessment of the problem-based learning tutorial session using a criterion-referenced system. *J Int Assoc Med Sci Educ*. 2004;14(1):8-14.
14. Sim SM, Azila NMA, Lian LH, Tan CPL, Tan NH. A simple instrument for the assessment of student performance in problem-based learning tutorials. *Ann Acad Med Singap*. 2006;35(9):634-41 [acesso em 20 jun 2019]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17051280/>.
15. Lee M, Wimmers PF. Validation of a performance assessment instrument in problem-based learning tutorials using two cohorts of medical students. *Adv Health Sci Educ Theory Pract*. 2016;21(2):341-57. doi: 10.1007/s10459-015-9632-y.
16. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manag*. 2004;42(1):15-29. doi: 10.1016/j.im.2003.11.002.
17. Chen JY, Lee MC, Lee HS, Wang YC, Lin LY, Yang JH. An online evaluation of problem based learning (PBL) in Chung Shan Medical University, Taiwan: a pilot study. *Ann Acad Med Sing*. 2006; 35(9): 624-33 [acesso em 20 jun 2019]. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17051279/>
18. Schuwirth LW, Van der Vleuten CPM. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach*. 2011;33(6):478-85. doi: 10.3109/0142159X.2011.565828.
19. Sluijsman DMA, Moerkerke G, Merriënboer JJGV, Dochy FJRC. Peer assessment in problem based learning. *Stud Educ Eval*. 2001;27(2):153-73 [acesso em 28 jun 2018]. Disponível em: https://sluijsmans.net/wp-content/uploads/2019/01/Sluijsmans-Peer-Assesment-in-Problem-Based-Learning-2001_9.pdf.
20. Lynch DC, Surdyk PM, Eiser AR. Assessing professionalism: a review of literature. *Med Teach*. 2004;26(4):366-73. doi: 10.1080/01421590410001696434.



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.