

Adaptação Transcultural da Escala de Autoeficácia de Realização Médica

Cross-Cultural Adaptation of the Medical Achievement Self-Efficacy Scale

Rebeca Carvalho Bressa¹, José Aloyseo Bzuneck², Leonardo de Oliveira Barros³,
Camélia Santana Murgo⁴ e Andressa Pereira de Souza⁵

Resumo

A autoeficácia acadêmica é um construto central na educação médica, influenciando estados emocionais, engajamento e estratégias de aprendizagem. Este estudo teve como objetivo adaptar e validar a Escala de Autoeficácia de Realização Médica (MASS) para o contexto brasileiro. Participaram 378 estudantes de medicina (69% mulheres; média de 24 anos; DP = 5,964). A análise da estrutura interna indicou uma solução unidimensional, integrando os diferentes domínios da competência médica em um único fator de autoeficácia. A regressão linear múltipla revelou que o semestre cursado e o tipo de ensino influenciam significativamente as pontuações da MASS. Os índices de validade de conteúdo e concordância entre juízes (CVC=.881–.929; $k=.488$) e participantes ($k=.757$) demonstraram a adequação da adaptação. Os resultados evidenciam que a versão brasileira da MASS é um instrumento válido e confiável para pesquisas e intervenções voltadas à formação médica.

Palavras-chave: autoeficácia, educação médica, validação, psicometria, estudantes de medicina

Abstract

Academic self-efficacy is a central construct in medical education, influencing emotional states, engagement, and learning strategies. This study aimed to adapt and validate the Medical Achievement Self-Efficacy Scale (MASS) for the Brazilian context. A total of 378 medical students participated (69% women; mean age=24 years; SD=5.964). Internal structure analysis supported a unidimensional solution, integrating different domains of medical competence into a single self-efficacy factor. Multiple linear regression indicated that academic semester and type of instruction significantly influenced MASS scores. Content validity indices and agreement among judges (CVC=0.881–.929; $k=.488$) and participants ($k=.757$) confirmed the adequacy of the adaptation. The findings demonstrate that the Brazilian version of the MASS is a valid and reliable instrument for research and educational interventions in medical training.

Keywords: self-efficacy, medical education, validation, psychometrics, medical students

¹Doutora em Educação. Universidade do Oeste Paulista (UNOESTE). Presidente Prudente, SP, Brasil. E-mail: rebecabressa@gmail.com ORCID: 0000-0001-5874-4502

²Doutor em Psicologia Escolar e do Desenvolvimento Humani pela Universidade de São Paulo (USP). Universidade Estadual de Londrina (UEL), Londrina, PR, Brasil. E-mail: bzuneck35@gmail.com ORCID: 0000-0003-3641-8284

³Doutor em Psicologia pela Universidade São Francisco. Universidade Federal da Bahia (UFBA), Salvador, BA, Brasil. E-mail: leonardobarros_lob@hotmail.com ORCID: 0000-0002-8406-0515

⁴Doutora em Psicologia pela Pontifícia Universidade Católica de Campinas (PUC-Campinas). Universidade do Oeste Paulista (UNOESTE), Presidente Prudente, SP, Brasil. E-mail: camelia@unoeste.br ORCID: 0000-0003-3932-7580

⁵Doutora em Educação pela Universidade do Oeste Paulista. Universidade do Oeste Paulista (UNOESTE), Presidente Prudente, SP, Brasil. E-mail: andressasouzap4@gmail.com ORCID:0000-0003-3730-2056

Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica. RIDEP · Nº79 · Vol. 1 · 165-179 · 2026

ISSN: 1135-3848 print /2183-6051online

This work is licensed under CC BY-NC 4.0. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>

Introdução

A formação médica é uma preocupação constante na sociedade, especialmente diante de transformações nas demandas em saúde e nas expectativas em relação ao papel do médico (Ferreira et al., 2023). Desde o relatório de Flexner (1910), consolidou-se um novo paradigma no ensino médico, orientado por uma visão mais científica e estruturada do ensino. Nas últimas décadas, a sociedade passou a demandar por profissionais envolvidos com a saúde coletiva, prevenção e redução da dependência hospitalar (Carraccio et al., 2002; Machado et al., 2018; Rooney, 2013; Silva & Baracat, 2016). Nesse cenário, destaca-se a adoção de currículos baseados em competências (CBC) com a intenção de articular o conhecimento essencial e as expectativas sobre as competências e atitudes dos futuros profissionais (Silva; Baracat, 2016).

O ensino por competências requer uma organização curricular que equilibre e alterne a aquisição de conhecimento com o desenvolvimento das habilidades e atitudes necessárias para a prática médica, buscando incluir todos esses domínios do aprendizado (Santos, 2011). Um dos CBC mais utilizados e reconhecidos na educação médica é o CanMEDS Framework (*Canadian Medical Education Directions for Specialists*). Este modelo formativo define sete domínios e 28 competências chaves necessárias para que os profissionais médicos atendam de forma eficiente às necessidades de saúde da população (Frank et al., 2015). A implementação de CBC exige, além de inovações curriculares, uma compreensão mais profunda dos fatores psicológicos que influenciam o processo formativo, entre eles, a autoeficácia.

A autoeficácia (AE) refere-se à crença do indivíduo em sua capacidade de organizar e executar ações necessárias para alcançar determinados objetivos (Bandura, 1997). No contexto educacional, as crenças de autoeficácia referem-se à percepção sobre sua capacidade de aprender, influenciando suas ações acadêmicas, escolhas curriculares, metas a seguir e perseverança diante das dificuldades relacionadas ao processo de aprendizagem (Bandura, 1993; Bzuneck, 2001; Schunk, 1995). A autoeficácia é crucial para sua integração no ensino superior

(Casanova & Polydoro, 2011). Estudantes com maior autoeficácia têm mais chances de sucesso, pois usam e testam suas habilidades com facilidade, são resilientes, aceitam desafios e superam falhas rapidamente. A crença aliada à competência torna-se um mecanismo poderoso de influência pessoal (Bandura, 1986, 1997; Bandura et al., 2008; Pajares & Olaz, 2008; Lopes, 2020).

Uma revisão sistemática sobre AE na educação médica evidenciou uma relação positiva entre a autoeficácia e diversos fatores associados ao desempenho acadêmico e ao bem-estar emocional de estudantes de Medicina, bem como à eficácia das práticas docentes (Costa Filho et al., 2022). Essas relações têm motivado o desenvolvimento de instrumentos específicos para avaliação da autoeficácia no contexto médico, como a *Medical Achievement Self-efficacy Scale* (MASS), elaborada por Turan et al. (2013). A MASS teve sua origem a partir da investigação acerca das capacidades de estudantes de medicina para atender às expectativas profissionais no currículo médico. A proposta da escala é avaliar a autoeficácia no domínio médico e sua relação com desenvolvimento de carreira e estado emocional. A MASS é unidimensional abrange todos os domínios de competências definidas na CanMEDS e apresentou boas propriedades psicométricas e alta confiabilidade ($\alpha=.89$) em sua versão original. Na amostra analisada, os escores de autoeficácia aumentaram significativamente ao longo dos anos. A análise de regressão mostrou que as pontuações da MASS previram significativamente as pontuações do teste de progresso de Maastricht.

No contexto brasileiro, estudos como os de Ferreira et al. (2025) e Kubrusly et al. (2024) utilizaram a Escala de Autoeficácia na Formação Superior (EAEFS) para avaliar a autoeficácia de estudantes e residentes em Medicina. Embora a EAEFS tenha apresentado adequação geral, sua estrutura fatorial revelou sete domínios específicos, refletindo as particularidades da formação médica. Essa limitação indica a necessidade de instrumentos mais sensíveis às especificidades do currículo médico e alinhados a modelos internacionais de competência profissional como a CanMEDS.

Dessa forma, destaca-se a importância de adaptar e validar, para o contexto brasileiro, uma medida específica de autoeficácia na educação

médica que contemple todos os domínios de competência relevantes. Considerando os resultados anteriores que demonstram a pertinência do modelo por sua abrangência e fundamentação teórica, a MASS representa uma alternativa promissora (Turan et al., 2013). Assim, com o intuito de contribuir para o avanço das práticas avaliativas e formativas na educação médica nacional, o objetivo deste estudo é adaptar e validar a Escala de Autoeficácia de Realização Médica (MASS) para o contexto brasileiro.

Método

Este estudo, de abordagem híbrida e transversal, foi desenvolvido na modalidade exploratória. Para atingir o objetivo proposto, o estudo foi organizado em quatro etapas: Tradução e *back-translation* (tradução reversa) da versão em inglês da MASS; Estimativas do Coeficiente de Validade de Conteúdo a partir da análise de juízes e a análise semântica; Estudo Piloto; Estudo de validação da estrutura interna da escala por meio da Análise Fatorial Exploratória e Confirmatória. Estas etapas são essenciais para estudos que avaliam as propriedades psicométricas dos instrumentos e sua validação apropriada para o contexto avaliado (Aragón, 2023; Luna et al., 2023; Santos et al., 2023; Uriarte et al., 2024).

Etapa 1 - Processo de tradução e retrotradução

Para traduzir e adaptar a Escala MASS, foram usadas as diretrizes de Beaton et al. (2000) com seis fases: tradução, síntese, retrotradução, comitê de revisão, pré-teste (estudo piloto) e submissão aos autores. Estas recomendações buscam equivalência semântica, conceitual e idiomática, corrigindo problemas gramaticais ou coloquiais.

A fase 1 (tradução) foi realizada de forma independente por dois professores da língua inglesa, sendo um nativo brasileiro e outro nativo inglês, mas fluente em português. O professor nativo brasileiro tinha domínio da Teoria Social Cognitiva (TSC), referencial que conceitualiza a autoeficácia. Essa seleção seguiu as recomendações de Beaton et al. (2000) que sugere que os dois tradutores devem ter perfis diferentes, um dos tradutores deve estar ciente dos conceitos que estão sendo examinados e o outro tradutor não deve ser consciente nem informado dos conceitos

que estão sendo quantificados e preferencialmente não deve ter antecedentes médicos ou clínicos. Isso é chamado de tradutor ingênuo.

Na etapa 2 (síntese), a pesquisadora contribuiu para a formulação da primeira versão traduzida como uma observadora e organizadora. Seguindo a recomendação de Beaton et al. (2000), os dois tradutores e um observador sentam-se para sintetizar os resultados das traduções. Trabalhando o questionário original, bem como o primeiro tradutor e as versões do segundo tradutor, uma síntese dessas traduções é realizada produzindo uma tradução comum (Beaton et al., 2000; 2007).

Na terceira fase (retrotradução), outros dois tradutores (bilíngues, nativo inglês e nativo brasileiro) realizaram a retro tradução (da escala), de maneira independente e sem contato prévio com a escala original. Como recomendado por Beaton et al. (2000), um terceiro e quarto tradutor, trabalhando com a versão do questionário e totalmente cego para a versão original, traduzem o questionário de volta para o idioma original. (Leplège & Verdier, 1995; Beaton et al., 2007).

A fase 4 (comitê de revisão / avaliação dos experts) foi realizada com análise da versão revisada pelos juízes especialistas nas áreas de autoeficácia, psicologia e educação médica. O papel do comitê de especialistas é consolidar todos as versões do questionário e desenvolver o que seria considerada a versão pré-final do questionário para o teste de campo. (Beaton et al., 2000; 2007).

Na quinta etapa (pré-teste) foi realizado estudo piloto, também descrito abaixo (Estudo Piloto) para verificar a concordância de uma amostra da população alvo da pesquisa. Idealmente, devem ser testadas entre 30 e 40 pessoas. Cada sujeito preenche o questionário / formulário avaliativo para investigar o que ele ou ela acha decida item da escala e a resposta escolhida. Isso garante que a versão adaptada ainda mantém sua equivalência em uma situação aplicada. (Beaton et al., 2000; 2007).

A sexta (submissão dos documentos aos autores) e, última etapa do processo de tradução, com definição final do instrumento em análise, foi realizada após contato por *e-mail* com a principal autora. A mesma concordou com a versão final retraduzida. Conforme Beaton et al. (2000; 2007) esta fase não deve alterar o conteúdo da versão final. É um processo com objetivo de garantir que

uma tradução razoável tem sido alcançada (Beaton et al., 2000; 2007).

Etapa 2 - Processo de análise de juízes

Nesta fase que teve como objetivo realizar a análise teórica da MASS, foram convidados via *e-mail* quatro pesquisadores *experts* com vasta experiência nas referidas temáticas (dois conhecedores da autoeficácia acadêmica e avaliação psicológica e dois específicos para a educação médica), respeitando o número mínimo de três e máximo de cinco juízes por temática proposto por Hernandez-Nieto (2002). A fim de analisar a viabilidade teórica do instrumento, foi estimado o coeficiente de validade de conteúdo (CVC), que consiste na mensuração da concordância entre os juízes em relação a cada item elaborado, ampliando a avaliação métrica comumente realizada pelo teste Kappa (Hernandez-Nieto, 2002; Pasquali, 2013).

Participantes do processo de análise de juízes

Os juízes que compuseram a amostra (N=4) foram dois *experts* do sexo feminino representando expertise em educação médica e dois *experts* do sexo masculino direcionados a AE acadêmica e avaliação psicológica/psicometria. Os especialistas em autoeficácia possuem doutorado e pós-doutorado em Psicologia e eram docentes de universidades públicas, da região Nordeste e Sul do Brasil. As *experts* em educação médica são docentes efetivas de Medicina em universidade privada na região Sudeste do Brasil, com doutorado em ciências cirúrgica e dos alimentos.

Instrumento

O roteiro avaliativo da MASS, instruiu os juízes a utilizarem uma escala Likert de 5 pontos para indicar o quanto o item estava adequado (1=pouquíssimo a 5=muitíssimo) diante das variáveis: 1) “Clareza de Linguagem” (CL), com objetivo de verificar o quanto a linguagem utilizada na redação dos itens estava suficientemente clara, compreensível e adequada para os futuros respondedores; 2) “Pertinência Prática” (PP), buscando analisar se cada item avaliava a autoeficácia acadêmica, a competência médica curricular e sua importância para o instrumento; 3) “Relevância Teórica” (RT), cujo objetivo é verificar o grau de concordância entre o item elaborado e o

referencial teórico adotado, verificando a existência de relação entre o item proposto e os construtos avaliados e; 4) “Dimensão teórica” (DT), verificando se o item é adequado ao fator para o qual foi elaborado, compreendendo a classificação de resposta “sim” ou “não” pelos juízes. Sugestões e comentários poderiam ser realizados no campo “Observações” ao lado de cada item.

Procedimentos de coleta e análise dos dados

Após a aprovação deste estudo pelo Comitê de Ética em Pesquisa da Universidade do Oeste Paulista (protocolo CAAE nº 46321321.2.0000.5515), foram enviados por e-mail aos participantes que aceitaram participar do estudo, o Termo de Consentimento Livre e Esclarecido, a Escala de Autoeficácia de Realização Médica (MASS) juntamente do roteiro avaliativo para orientar a análise teórica dos itens (Pasquali, 2013).

Consecutivo a avaliação dos juízes, os dados foram inseridos numa planilha do Excel e analisados, sendo calculado o CVC para cada item do instrumento (CVCc) e para o instrumento como um todo (CVCt). Autores sugerem que itens de boa qualidade devem apresentar valores maiores que .80 (Hernandez-Nieto, 2002; Balbinotti et al., 2007; Pasquali, 2013). Os itens que não atingirem tais percentuais podem ser reformulados apenas nas dimensões “Relevância Teórica” e “Clareza da Linguagem” anteriormente a coleta de dados (Greco et al., 2014).

Resultados da análise de juízes

A análise de concordância foi realizada através do Kappa de Fleiss e do Coeficiente de Validade de Conteúdo (Landis & Koch, 1977; Hernandez-Nieto, 2002). Os resultados gerais indicaram bons níveis de concordância entre os juízes em relação à clareza (CL), pertinência (PP), relevância (RT) e dimensão (DT), conforme explicitados pelas Tabelas 1 e 2, onde os valores de CVCt variaram entre .881 e .929. A concordância pelos *experts* foi moderada pelo teste de Kappa de Fleiss ($k=.488$, $p<.05$) (Landis & Koch, 1977). Após a análise dos juízes a escala foi mantida e seguiu para a fase do estudo piloto sem alterações significativas e sem exclusão de qualquer item.

Etapa 3 - Estudo piloto

Para esta etapa, foram convidados 31 estudantes de medicina de uma universidade privada do interior de São Paulo, cursando diferentes períodos da graduação. Todos assinaram o termo de consentimento para participar da pesquisa. Cabe salientar, que se tratou de uma amostra por conveniência e não aleatória, sendo intencional que houvesse participantes heterogêneos. A amostra foi composta por: alunos do primeiro semestre ($n=10$, 32.3%) e estudantes do último semestre ($n=21$, 67.7%), sendo a maioria do sexo feminino ($n=18$, 58.1%), seguida por participantes do sexo masculino ($n=13$, 41.9%). A idade mínima do grupo foi de 18 anos e a máxima 33 anos, sendo registradas as seguintes frequências para os intervalos: entre 18 a 21 anos ($n=6$, 19.3%), entre 22 a 25 anos ($n=12$, 38.7%), entre 26 a 30 anos ($n=10$, 32.3%) e entre 31 a 33 anos ($n=3$, 9.7%).

Instrumentos do Estudo Piloto

Questionário para avaliação da clareza dos itens da versão adaptada da escala MASS: composto de perguntas relacionadas ao grau de compreensão da redação dos itens e considerações diversas que pudessem aprimorar a semântica e clareza da MASS na língua portuguesa para estudantes de medicina, com as seguintes questões: 1 - Você conseguiu entender as instruções para responder a escala?; 2 - Há alguma palavra na instrução que você não conhece o significado?; 3 - Você achou alguma parte da escala confusa?; 4 - Você encontrou itens (frases) com conteúdo muito parecido?. O aluno tinha espaço para escrever sua opinião e deveria responder entre “sim” e “não”.

Procedimento de coleta e análise de dados

A coleta do estudo piloto foi executada por meio da escala MASS versão T-12 (já avaliada pelos juízes) impressa acompanhada de um formulário avaliativo também impresso e entregue pessoalmente para cada participante, com intuito de realizar uma análise semântica do instrumento (Pasquali, 2013).

A análise de concordância foi realizada através do Kappa de Fleiss (Landis & Koch, 1977). Entre o público-alvo a concordância é substancial ($k=.757$, $p<.05$) (Landis & Koch, 1977). Quanto aos aspectos da avaliação do público-alvo,

respondidas por itens dicotômicos (sim ou não). Com as seguintes perguntas: 1- Você conseguiu entender as instruções para responder a escala?; 2- Há alguma palavra na instrução que você não conhece o significado?; 3- Você achou alguma parte da escala confusa?; 4- Algum item não compreensível?; 5- Você encontrou itens (frases) com conteúdo muito parecido?

Os participantes (100%) não encontraram dificuldades de compreensão da redação e linguagem nas instruções para responder a escala. Apenas quatro (12.9%) alunos referiram não conhecer a palavra ortotanásia, sendo esses todos (100%) estudantes do primeiro período da graduação e três (9.7%) apontaram o item 6 (aspectos médicos nas ciências humanas) como confuso. Esse fato coincidiu com estudo piloto original realizado por Turan et al. (2013) e provavelmente é explicado pelo mesmo motivo: ausência desse assunto no currículo médico a qual os estudantes eram vinculados. Apesar disso, em concordância com a relevância do item 6 para as competências CanMeds, já descrito por Turan et al. (2013), este item foi mantido.

Face ao exposto, apenas foi acrescentado entre parênteses o significado da palavra ortotanásia (morte natural, sem antecipação ou prolongamento) no item 11 da escala. Nenhum dos itens foi excluído, encerrando, assim, o processo de adaptação e tradução da escala MASS para a língua portuguesa.

Etapa 4 - Estudo de validação da estrutura interna da escala MASS

Para a validação da estrutura interna é mais comum utilizar a técnica de Análise Fatorial Exploratória (AFE) e análise de consistência interna dos itens. A busca por evidências compreende a verificação da coesão da estrutura prevista pelo construto, com a observada na experimentação. Trata-se de uma técnica que procura descobrir as variáveis latentes, capazes de explicar o comportamento observado, a partir de análise estatística. Devido as limitações teóricas da AFE, a Análise Fatorial Confirmatória (AFC) tem sido mais usada, pois o modelo fatorial é mais adequado e pode ser avaliado por diferentes índices de ajuste (Fulmer et al., 2014; Peixoto & Ferreira-Rodrigues, 2019).

Participantes

A amostra contou com 378 estudantes brasileiros, sendo 79,8% (n=302) de universidades privadas, com idade média de 24.077 anos (DP=5.964), sendo 261 (69.0%) do sexo feminino. A maior parte dos estudantes não tinha filhos (92%; n=348), não possuía outra graduação (84.6%; n=320) e não exercia atividade remunerada (84.9%; n=321). Entre os participantes, apenas 65 (17.19%) exerciam atividade de trabalho relacionada com a faculdade. Por fim, 201 (53.1%) fazem alguma atividade extracurricular.

Instrumentos

Questionário Sociodemográfico de Caracterização da Amostra (QSCA): elaborado para verificar as características sociodemográficas dos participantes da pesquisa. Contém 16 (dezesesseis) questões, sendo as variáveis analisadas: 1. Nome; 2. Idade; 3. Sexo; 4. Estado Civil; 5. Nacionalidade; 6. Possui Filhos; 7. Região do Brasil em que mora; 8. Com quem vive atualmente; 9. Tipo de Instituição; 10. Metodologia de Ensino da universidade que estuda; 11. Possui outro curso de graduação; 12. Utiliza-se de algum destes recursos para estar na Universidade; 13. Semestre atual do curso; 14. Exerce atividade remunerada; 15. Sua atividade laboral está relacionada com seu curso; 16. Exerceu ou exerce alguma atividade extracurricular.

Escala MASS (Turan et al., 2013 adaptada para este estudo): tem como objetivo verificar as crenças dos estudantes de Medicina no desenvolvimento das competências profissionais cognitivas, psicomotoras e afetivas à medida que avançam o currículo médico. A MASS utiliza a estrutura de competências CanMEDS e é composta por 18 itens que medem as competências do médico em sete domínios: 1 - Especialista (itens 1, 4, 7); 2 - Comunicador (itens 5, 15); 3 - Colaborador (itens 16, 12); 4 - Líder (itens 8, 18); 5 - Advogado da Saúde (itens 2, 14); 6 - Acadêmico (itens 3, 6, 9, 10) e 7 - Profissional (itens 11, 12, 17). A escala é respondida em uma escala Likert de 5 pontos, variando de 1 -pouco a 5 -muito.

Procedimento de coleta de dados e análise de dados

A recolha ocorreu através da plataforma *Google Forms*, sendo sua divulgação por meio

digital. A plataforma *Google Forms* continha os seguintes documentos: TCLE (Termo de Compromisso, Livre e Esclarecido), Escala de autoeficácia MASS (versão adaptada final) e Questionário Sociodemográfico. Os participantes responderam à pesquisa por meio de um *link* e, só puderam prosseguir as respostas após a aceitação do TCLE.

O estudo se divide em três grandes análises: (1) Análise Fatorial Exploratória (AFE), (2) Análise Fatorial Confirmatória (AFC) e (3) Regressão Linear Múltipla (RLM). Após a coleta de dados foi realizado a divisão da amostra, de forma aleatória pelo *Software Excel*, em dois grupos (grupo 1 e 2), cada um para cada análise, AFE e AFC.

A AFE (análise 1) foi feita com o objetivo de compreender a estrutura fatorial da MASS. Realizada no *FACTOR*. A matriz de correlação utilizada foi a policórica associado ao *bootstrap* (*Bias-corrected and accelerated*, 500 re-amostragens) (Lambert et al., 1991) e ao *Robust Diagonally Weighted Least Squares* (RDWLS), a plausibilidade da matriz foi avaliada pelo teste de *Kaiser-Meyer-Olkin* (KMO) e a esfericidade de Bartlett (Distefano & Morgan, 2014). A retenção fatorial foi realizada pela Análise Paralela com a substituição aleatória dos dados coletados (Timmerman & Lorenzo-Seva, 2011) e a rotação utilizada foi a *Robust Promin* (Lambert et al., 1991; Ferrando & Lorenzo-Seva, 2018).

Os índices de ajuste utilizados foram: Root Mean Square Error of Aproximation (RMSEA), Comparative Fit Index (CFI) e Tucker-Lewis Index (TLI) (Brown, 2006). Foi utilizado o Índice *H* e a fidedignidade composta para avaliar o quão bem o conjunto de itens representavam o fator (Ferrando & Lorenzo-Seva, 2018; Hancock & Mueller, 2001). Por fim, foi utilizado os thresholds dos itens com a parametrização de Reckase (Reckase, 1985). Threshold é um parâmetro da TRI que mostra a quantidade de traço latente que é necessário para que haja mudança no padrão de resposta, como a escala é ordinal se espera que esses valores sejam ascendentes (Reckase, 1985).

A análise fatorial confirmatória (análise 2) foi realizada a fim de avaliar a plausibilidade da estrutura proposta na análise fatorial exploratória da MASS no software JASP. A análise teve como método de estimação Robust Diagonally Weighted Least Squares (RDWLS), adequado para dados

Tabela 1. Resultados da análise fatorial exploratória e *thresholds*

Retenção fatorial, Análise paralela				
		Percentual de variância explicada dos dados reais	Percentual de variância explicada média dos dados aleatórios	Percentual de variância explicada dos dados aleatórios (95% IC)
Número de fatores				
1		52.122	11.575	12.8
2		8.038	10.487	11.506
Carga fatorial, matriz não rotacionada				
Item		Carga fatorial		
1		.732		
2		.616		
3		.488		
4		.769		
5		.693		
6		.57		
7		.764		
8		.748		
9		.802		
10		.758		
11		.598		
12		.644		
13		.507		
14		.742		
15		.816		
16		.753		
17		.646		
18		.709		
Threshold				
		Categoria 1-2	Categoria 2 - 3	Categoria 3 - 4
Item				Categoria 4 - 5
1		-1.747	-.838	.357
2		-3.295	-1.893	-.249
3		-4.402	-2.621	-.883
4		-1.329	-.269	1.065
5		-2.202	-1.38	-.105
6		-2.003	-.834	.731
7		-1.996	-.977	.078
8		-1.492	-.518	.595
9		-.976	-.361	.725
10		-1.008	-.096	1.105
11		-1.53	-.6	.671
12		-3.004	-2.032	-.903
13		-3.401	-1.579	.197
14		-2.406	-1.377	-.063
15		-2.045	-1.074	-.122
16		-2.086	-1.164	-.132
17		-1.515	-.513	.448
18		-1.41	-.507	.608
Fidedignidade composta				
H-Latent				.942
H-Observed				1.006
Qualidade e eficácia das estimativas de pontuação dos fatores				
Factor Determinacy Index (FDI)				.974
Orion				.949
Sensitivity ratio (SR)				4.136
Expected percentage of true differences (EPTD)				95.20%
Confiabilidade				
ômega de McDonald				.936
Alfa de Cronbach				.936
Greatest Lower Bownd				.978

Fonte: A autora.

categóricos (Distefano & Morgan, 2014). Os índices de ajustes avaliados foram o CFI, TLI e RMSEA. Valores de CFI e TLI devem ser > que .90 e, preferencialmente acima de .95; Valores de RMSEA devem ser < que .08 ou, preferencialmente < que .06, com intervalo de confiança (limite superior) <.10 (Brown, 2015). Por fim, foi realizada uma regressão linear múltipla, análise 3, com o objetivo de avaliar os impactos do semestre em que

o aluno estava junto do tipo de ensino (PBL, Híbrido ou tradicional) nos níveis da MASS. O método utilizado foi o enter implementado do Bootstrap (bias corrected and accelerated 100 reamostragens), a fim de corrigir desvios de distribuição. “A variável tipo de Ensino” foi transformada em uma variável dummy, tendo como referência o ABP (Field, 2012).

Resultados do estudo de validade interna

Etapa 1: Análise Fatorial Exploratória

O teste de esfericidade de Bartlett (2085.8, $gl=153$, $p<.001$) e de Kaiser-Meyer-Olkin (KMO=.868, IC 90%=0.739-0.868) apontaram fatorabilidade da matriz e a análise paralela sugeriu um modelo de um fator uma vez que o percentual da variância explicada é maior que os dados randômicos, tal como exposto na Tabela 1. Como se esperava uma estrutura fatorial semelhante ao do estudo original foi verificada. Três índices indicaram a unidimensionalidade, corroborando os resultados da análise paralela: UniCo=.962 (IC 95%=.946-.984), ECV=.875 (IC 95%=.865-.911) e MIREAL=.180 (IC 95%=.163-.173) (Ferrando & Lorenzo-Seva, 2018). Ver Tabela 1.

Os índices de ajuste CFI (0,968) e TLI (0,063) demonstram que o modelo unidimensional é plausível, já o RMSEA (.104) é próximo do limite máximo aceitável, (<.100). No grupo piloto do estudo original da escala MASS (n=94), as correlações item-total variaram de .33 a .61 ($p<.001$); com exceção do item sobre “humanidades médicas” (item 6) ($r=-.10$). O coeficiente de α de Cronbach foi de .85 no estudo piloto. Esses resultados já podiam ser usados como um primeiro indicador sobre a validade discriminante da escala. Todos os itens discriminaram significativamente entre alunos de baixo e alto desempenho ($p<.001$); com exceção do item 6. Este resultado foi justificado pelos autores devido à ausência de “humanidades médicas” no currículo médico de Leuven. Uma vez que este item é uma parte importante do médico competências do CanMEDS e do currículo Ghent, não foi excluído (Turan et al., 2013).

No estudo principal, as correlações item-total variaram entre .39 e .64 ($p<.001$). A comparação entre as pontuações dos grupos de baixo e alto desempenho mostrou que todos os itens discriminaram consistentemente entre os dois grupos de níveis de desempenho ($p<.001$). A consistência interna de toda a escala é alta; α de Cronbach=.89 (Turan et al., 2013).

Ao comparar os achados do estudo original e do Brasil, observa-se similaridade da confiabilidade da escala MASS com coeficiente α de Cronbach de .936. No estudo original, os autores utilizaram apenas a correlação de item-total. Já nesta pesquisa, no Brasil, houve ampliação com

uso da TRI (teoria de resposta ao item) utilizando Threshold conforme Tabela 4. A variável latente foi bem definida, acima de .8 (.942), o que é mais provável que seja estável em diferentes estudos (Ferrando & Lorenzo-Seva, 2018).

Etapa 2: Análise Fatorial Confirmatória

Na AFC da MASS os índices de ajustes foram contraditórios, O CFI e TLI obtiveram valores satisfatório, enquanto o RMSEA não (Tabela 2, Figura 1), assim foram observados os índices de ajustes.

Tabela 2. Índices de ajuste da MASS

Modelo	CFI	TLI	RMSEA (90% IC)
Original	.988	.986	.110 (.098 - .121)
Reespecificado	.999	.999	.027 (.00 - .047)

Fonte: A autora.

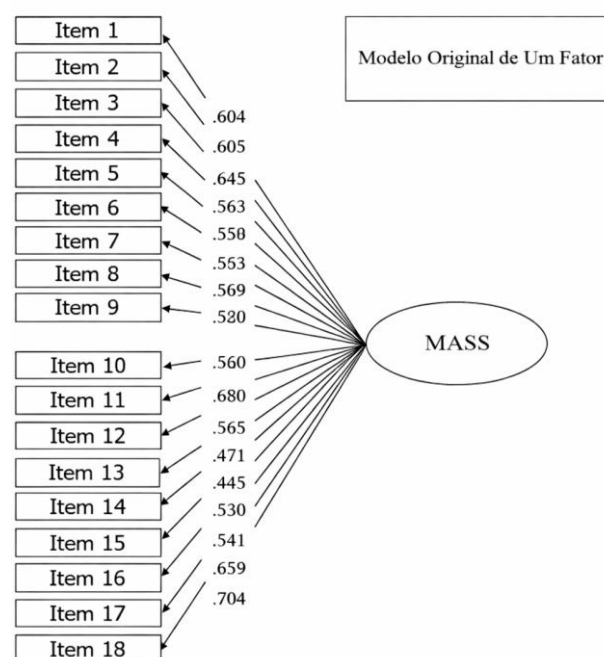


Figura 1. Índices de ajuste da (MASS)

Fonte: A autora.

Os itens 9 e 10 apresentaram alto valor de índice de modificação (288.493), assim optado pela retirada do item 10, o qual levou a um modelo mais parcimonioso (Tabela 2, Figura 2).

Etapa 3: Regressão Linear Múltipla

A regressão linear é uma técnica utilizada para compreender relações entre variáveis. Seu principal objetivo é obter uma equação que explique a relação entre uma variável dependente e

Tabela 3. Coeficientes da regressão linear múltipla

Variáveis	B	Erro Padrão	B padronizado	Intervalo de Confiança (95%)		t	p	VIF
				Limite inferior	Limite Superior			
Intercepto	52.764			49.162	56.366	28.804		
Semestre	1.517	1.832	.318	1.062	1.972	6.559	<.001	1.003
Metodologia de Ensino Híbrido	-6.355	.231		-9.662	-3.049	-3.78	<.001	1.002
Metodologia de Ensino Tradicional	-7.535	1.681		-11.561	-3.509	-3.68	<.001	

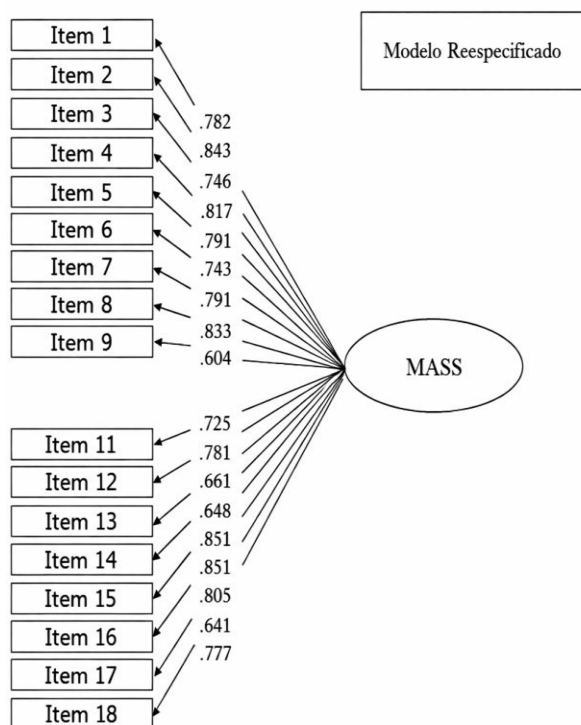


Figura 2. Índices de ajuste da (MASS)

Fonte: A autora.

outra (s) variáveis independentes explicativas, possibilitando fazer predição e comparações. Logo com esse modelo é possível prever tendências ou comportamentos. Na regressão linear múltipla pode haver mais de uma variável preditiva (independente) do tipo ordinal, escalar ou categórica (Gonçalves & Silva et al., 2023).

No estudo original da escala MASS, Turan et al. (2013) levantaram a hipótese que a AE evoluiria ao longo dos anos do currículo médico e realizaram uma análise fatorial unidirecional para comparar a pontuação da autoeficácia de acordo com o ano curricular. Houve significância estatística e, como esperado, as pontuações de AE aumentaram ao longo dos anos (Turan et al., 2013). Optou-se por fazer investigação comparativa com os dados brasileiros e a regressão linear múltipla demonstrou que ambas as variáveis (semestre e tipo de ensino) tiveram um impacto

estatisticamente significativo nos níveis da MASS [F (3, 373)=18,868, $p<.001$; $R^2=.125$; Durbin Watson=2,065, $p=.556$]. Quanto mais avançado o semestre, maiores foram os níveis da MASS (B1,517; B padronizado=.318) e a metodologia de ensino ABP apresentou maiores níveis (6,355 pontos a mais que ensino híbrido e 7,535 a mais que o tradicional), ver Tabela 3.

Discussão

Esta pesquisa traduziu, adaptou e avaliou a Escala de Autoeficácia para Realização Médica (MASS) para o Brasil. As propriedades psicométricas observadas indicam que o instrumento traduzido apresenta evidências de validade com base no conteúdo do teste, na estrutura interna e na relação com variáveis externas ao teste, além de confiabilidade para avaliar a autoeficácia dos estudantes de medicina. O processo de tradução e adaptação permitiu obter uma versão compreensível e pertinente ao contexto brasileiro, seguindo as recomendações da Fulmer et al. (2014). As avaliações realizadas por docentes e pesquisadores de Medicina e da mensuração em Psicologia, indicaram que o instrumento tem adequação teórica e consistência na formulação dos itens. Na mesma direção, o estudo piloto com a população-alvo demonstrou que as instruções, itens e chave de resposta são de fácil compreensão, aspecto que possibilita a aplicação da escala com estudantes de medicina.

A análise de dimensionalidade da escala indicou que a MASS avalia autoeficácia em uma única dimensão, apesar de abordar os diferentes domínios de competências médica da CanMEDS. A unidimensionalidade é desejável na avaliação fatorial, já que homogeneidade faz sentido ao avaliar um único construto (Vitória et al., 2006). Assim, a escala adaptada é confiável para uso em estudantes de medicina, pois o conceito de

autoeficácia mostrou-se uniforme e coeso. De tal modo, no presente estudo replicou-se a estrutura interna e foi obtido índice de precisão adequado, tal como no estudo original de Turan et al (2013).

A regressão linear múltipla indicou que tanto o semestre quanto o tipo de ensino tiveram um impacto estatisticamente significativo nos níveis da MASS do Brasil. Semestres mais avançados apresentaram maiores níveis da MASS, e a metodologia de ensino de Aprendizagem Baseada em Problemas (ABP) mostrou níveis mais altos. No estudo original de Turan et al. (2013), foi realizada uma análise de variância unidirecional para comparar a pontuação de autoeficácia dos estudantes de medicina conforme o ano curricular. As pontuações de autoeficácia diferiram significativamente, aumentando ao longo dos anos de estudo, assim como no presente estudo. Tais resultados sugerem que a exposição à situação práticas e aquisição gradativa de conhecimento tende a deixar os estudantes mais confiantes em relação às suas capacidades de executar tarefas próprias da formação médica (Lopes et al., 2020).

A autoeficácia é reconhecida como um dos principais preditores do desempenho individual, influenciando diretamente a motivação, a persistência diante de desafios e a qualidade da atuação em diferentes contextos. Conforme argumentam Klassen e Klassen (2018), a percepção de autoeficácia atua como um mecanismo motivacional fundamental, sendo frequentemente um pré-requisito para o engajamento efetivo em atividades de desempenho. No contexto da educação médica, Ferreira et al. (2025) destacam a relevância de mensurar e reavaliar periodicamente os níveis de autoeficácia dos estudantes como estratégia para identificar avanços e fragilidades no processo formativo. As fontes de autoeficácia (experiência pessoal, aprendizagem vicária, persuasão verbal e indicadores fisiológicos) podem ser intencionalmente cultivadas no ambiente universitário, por meio de práticas pedagógicas que estimulem a autonomia, o feedback construtivo e a participação ativa dos discentes (Guerreiro-Casanova et al., 2011; Bandura, 1986).

Diversos estudos têm contribuído para o avanço da literatura ao desenvolver, adaptar e utilizar escalas voltadas à mensuração da autoeficácia no contexto da educação médica

(Klassen & Klassen, 2018; Costa Filho et al., 2022). Um exemplo notável é o trabalho de Pololi e Price (2000), que criaram e validaram um instrumento composto por 31 itens com o objetivo de avaliar o ambiente de aprendizagem percebido por estudantes de medicina nos Estados Unidos. A pesquisa, investigou três dimensões principais: a relação professor-aluno (*Teacher-Learner Relationship* – TLR), a relação médico-paciente (*Physician-Patient Relationship* – Phys-Pt R) e a percepção de autoeficácia. Os resultados indicaram que, à medida que os estudantes avançavam no curso, havia uma redução nas percepções positivas em todas as dimensões avaliadas, incluindo a autoeficácia. Como resposta a esse achado, foram implementados programas de capacitação docente com foco no desenvolvimento de práticas pedagógicas que favorecessem a construção de vínculos educacionais e uma aprendizagem centrada no estudante. Embora o instrumento tenha demonstrado bons índices de confiabilidade (α entre .81 e .94), ele não contempla integralmente as demandas e a estrutura dos currículos médicos contemporâneos, especialmente aqueles orientados por competências.

Nesse contexto, a MASS se destaca por ter origem em um estudo mais amplo e voltado à avaliação de inovações curriculares na educação médica, especialmente no que se refere ao desenvolvimento de competências profissionais (Turan et al., 2013). Desde então, a escala tem sido aplicada em diferentes contextos e fases da formação médica, demonstrando sua relevância. Tawfeek et al. (2021), por exemplo, utilizaram a MASS para investigar a relação entre o papel do instrutor clínico, a autoeficácia dos estudantes e o desempenho acadêmico. Os autores identificaram uma correlação positiva significativa entre os níveis de autoeficácia (avaliados pela MASS), as qualidades do instrutor e os resultados acadêmicos, ressaltando o papel central do educador clínico como mentor no fortalecimento da confiança e da autonomia dos discentes. Em outra aplicação, Batur et al. (2024) analisaram a relação entre o fenômeno do impostor (FI) e a autoeficácia acadêmica, constatando uma correlação negativa robusta: quanto maior a intensidade do FI, menor a percepção de autoeficácia entre os estudantes. Esses achados reforçam a sensibilidade da MASS para captar variações psicológicas relevantes

durante a formação médica, além de seu potencial como ferramenta de rastreio e de apoio ao desenvolvimento acadêmico e emocional dos alunos.

Considerações Finais

Instrumentos que medem a AE acadêmica e que relacionam metodologias ativas e o progresso curricular podem fornecer informações essenciais para melhorar o envolvimento dos alunos e as bases curriculares. Considerando a escassez de instrumentos para avaliar as crenças dos estudantes de medicina no desenvolvimento acadêmico, a validação da escala MASS no Brasil é importante e pode auxiliar intervenções que aumentem a AE acadêmica durante a graduação médica. Assim, o presente estudo apresenta a escala MASS adaptada ao contexto brasileiro com boa qualidade psicométrica e com potenciais aplicações em instituições de ensino.

Este estudo representa um avanço importante ao adaptar e validar a Escala de Autoeficácia de Realização Médica (MASS) para o contexto brasileiro. No entanto, algumas limitações metodológicas e contextuais devem ser consideradas na interpretação dos resultados. Primeiramente, o delineamento transversal limita a compreensão da evolução da autoeficácia ao longo do tempo, restringindo inferências sobre causalidade. Embora os resultados indiquem diferenças nos níveis de autoeficácia de acordo com o semestre e a metodologia de ensino, não é possível afirmar com precisão como essas crenças se desenvolvem ao longo da formação médica. Estudos longitudinais são recomendados para acompanhar as trajetórias individuais de autoeficácia durante o curso de Medicina e identificar fatores contextuais ou pedagógicos que promovam ou inibam seu fortalecimento.

Em segundo lugar, a amostra utilizada foi composta predominantemente por estudantes de instituições privadas, o que pode limitar a generalização dos achados para o conjunto da população universitária brasileira, especialmente aqueles vinculados a universidades públicas. Pesquisas futuras devem buscar maior diversidade amostral, incluindo diferentes tipos de instituições, metodologias de ensino, regiões geográficas e perfis socioculturais dos estudantes. Outra limitação refere-se à ausência de análise de

validade convergente e discriminante com outros instrumentos já validados no contexto brasileiro. Embora a estrutura unidimensional da MASS tenha se mostrado robusta, a investigação da correlação com escalas que avaliem construtos relacionados, como motivação, estresse acadêmico, desempenho escolar ou bem-estar psicológico, pode fortalecer as evidências de validade da versão brasileira do instrumento. Por fim, recomenda-se investigar o uso da MASS como ferramenta diagnóstica e formativa em práticas pedagógicas específicas (e.g. tutorias, mentorias, feedback clínico e metodologias ativas) a fim de explorar seu potencial para monitoramento do desenvolvimento de competências e para intervenções pedagógicas voltadas ao fortalecimento da autoeficácia discente.

Referências

- Aragón, R. S. (2023). Validación de la versión Mexicana del Inventario de Estilos de Amor Aplicado al Duelo Romántico. *Revista Iberoamericana de Diagnóstico y Evaluación-e Avaliação Psicológica*, (70), 73-84.
<https://doi.org/10.21865/RIDEP70.4.06>
- Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs*, 1986(23-28), 2.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.
- Bandura, A., Azzi, R. G., & Polydoro, S. A. (2009). *Teoria social cognitiva: Conceitos básicos*. Artmed Editora.
- Alencar Abaide Balbinotti, M., Benetti, C., & Renato Soares Terra, P. (2007). Translation and validation of the Graham-Harvey survey for the Brazilian context. *International Journal of Managerial Finance*, 3(1), 26-48.
<https://doi.org/10.1108/17439130710721644>
- Batur, A., Arslan, V., & Turan, S. (2024). Analysis of the relationship between medical achievement self-efficacy perception and impostor phenomenon in the postgraduate education process: A cross-sectional study. *BMC Medical Education*, 24(1), 1536.
<https://doi.org/10.1186/s12909-024-06591-9>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.

- <https://doi.org/10.1097/00007632-200012150-00014>
- Beaton, D., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2007). Recommendations for the cross-cultural adaptation of the DASH & QuickDASH outcome measures. *Institute for Work & Health*, 1(1), 1-45.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (1st ed.). The Guilford Press.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Bzuneck, J. A. (2001). O esforço nas aprendizagens escolares: Mais do que um problema motivacional do aluno. *Revista Educação e Ensino*, 6(1), 7-18.
- Carraccio, C., Wolfsthal, S. D., Englander, R., Ferentz, K., & Martin, C. (2002). Shifting paradigms: From Flexner to competencies. *Academic Medicine*, 77(5), 361-367. <https://doi.org/10.1097/00001888-200205000-00003>
- Costa Filho, J. O., Murgo, C. S., & Franco, A. F. (2022). Autoeficácia na educação médica: Uma revisão sistemática da literatura. *Educação em Revista*, 38, e35900. <https://doi.org/10.1590/0102-469835900>
- DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 425-438. <https://doi.org/10.1080/10705511.2014.915373>
- Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, 78(5), 762-780. <https://doi.org/10.1177/0013164417719308>
- Ferreira, M. M. D. S., Maid, L. C., Costa, S. D. M., & Caldeira, A. P. (2023). Diretrizes curriculares nacionais para os cursos de medicina no Brasil: Mudanças no processo de formação. *Jornal de Políticas Educacionais*, 17, e89451. <https://doi.org/10.5380/jpe.v17i0.89451>
- Ferreira, M. R. C., Fernandes, R. A. F., Malloy-Diniz, L. F., & Ibiapina, C. (2025). Autoeficácia ou inteligência para predizer o sucesso acadêmico ou profissional?. *Revista Brasileira de Educação Médica*, 49(2), e057. <https://doi.org/10.1590/1981-5271v49.1-2023-0010>
- Field, A. (2012). *Discovering statistics using R*. SAGE Publications.
- Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the Advancement of Teaching*. Updyke.
- Frank, J. R., Snell, L., & Sherbino, J. (Eds.). (2015). *CanMEDS 2015: Physician competency framework*. Royal College of Physicians and Surgeons of Canada. https://canmeds.royalcollege.ca/uploads/en/framework/CanMEDS%202015%20Framework_EN_Reduced.pdf
- Fulmer, M. (2014). American Educational Research. *American Educational Research Journal*, 51(6), 1195-1226.
- Gonçalves e Silva, E., Vieira, C. E., & Polydoro, S. A. J. (2023). *Análises estatísticas com JASP: Um guia introdutório*. Creative Commons Attribution Non-Commercial 4.0 International. <https://jasp-stats.org/wp-content/uploads/2023/04/v.2-Analises-Estatisticas-com-JASP-Um-Guia-Introdutorio.pdf>
- Greco, P. J., Morales, L. M., & Pedrosa, R. H. (2014). Validação de conteúdo de ações tático-técnicas do Teste de Conhecimento Tático Processual - Orientação Esportiva. *Motricidade*, 10(1), 38-48. [https://doi.org/10.6063/motricidade.10\(1\).2124](https://doi.org/10.6063/motricidade.10(1).2124)
- Guerreiro-Casanova, D. C., & Polydoro, S. A. J. (2011). Autoeficácia na formação superior: Percepções durante o primeiro ano de graduação. *Psicologia: Ciência e Profissão*, 31(1), 50-65. <https://doi.org/10.1590/S1414-98932011000100006>
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. *Structural Equation Modeling*, Manuscript submitted for publication. https://www.researchgate.net/publication/312447691_Rethinking_construct_reliability_within_latent_variable_systems

- Hernández-Nieto, R. A. (2002). Contributions to statistical analysis. *Mérida: Universidad de Los Andes*, 193.
- Klassen, R. M., & Klassen, J. R. (2018). Self-efficacy beliefs of medical students: a critical review. *Perspectives on medical education*, 7, 76-82.
<https://doi.org/10.1007/s40037-018-0411-3>
- Kubrusly, M., de Aquino, B. O. A., Simonian, T. S., do Nascimento Oliveira, M., & Rocha, H. A. L. (2024). Self-efficacy of medical students in a hybrid curriculum course (traditional and problem-based learning) and associated factors. *BMC Medical Education*, 24(1), 9.
<https://doi.org/10.1186/s12909-023-05016-3>
- Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1991). Approximating confidence intervals for factor loadings. *Multivariate Behavioral Research*, 26(3), 421-434.
https://doi.org/10.1207/s15327906mbr2603_3
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
<https://doi.org/10.2307/2529310>
- Leplège, A. (1995). The adaptation of health status measures: A discussion of certain methodological aspects of the translation procedure. *The international assessment of health-related quality of life: Theory, translation, measurement and analysis*.
- Lopes, J. M., Castro, J. G. F., Peixoto, J. M., & Moura, E. P. (2020). Autoeficácia de estudantes de medicina em duas escolas com metodologias de ensino diferentes (aprendizado baseado em problemas versus tradicional). *Revista Brasileira de Educação Médica*, 44, e047.
<https://doi.org/10.1590/1981-5271v44.2-20190187>
- Luna, S. M. A., Rodríguez, O. F., de la Torre, A. E. H., Garduño, M. D. L. V., & Betanzos, F. G. (2023). Adaptación, Validación y Propiedades Psicométricas de la DBT-WCCL-DSS en Población Mexicana. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 4(70), 175-185.
<https://doi.org/10.21865/RIDEP70.4.13>
- Machado, C. D. B., Wuo, A., & Heinzle, M. (2018). Educação médica no Brasil: Uma análise histórica sobre a formação acadêmica e pedagógica. *Revista Brasileira de Educação Médica*, 42(4), 66-73.
<https://doi.org/10.1590/198152712015v42n4R-B20180065>
- Pajares, F., & Olaz, F. (2008). Teoria social cognitiva e auto-eficácia: Uma visão geral. *Teoria social cognitiva: conceitos básicos. Porto Alegre: Artmed*, 97, 114.
- Pasquali, L. (2013). *Psicometria: Teoria dos testes na psicologia e na educação*. Vozes.
- Peixoto, E. M., & Ferreira-Rodrigues, C. F. (2019). Propriedades psicométricas dos testes psicológicos. *Compêndio de avaliação psicológica*, 29-39.
- Pololi, L., & Price, J. (2000). Validation and use of an instrument to measure the learning environment as perceived by medical students. *Teaching and learning in Medicine*, 12(4), 201-207.
- Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9(4), 401-412.
<https://doi.org/10.1177/014662168500900409>
- Rooney, A. (2013) *A história da medicina: das primeiras curas aos milagres da medicina moderna*. M. Books do Brasil Editora.
- Santos, W. S. D. (2011). Organização curricular baseada em competência na educação médica. *Revista Brasileira de Educação Médica*, 35, 86-92.
<https://doi.org/10.1590/S0100-55022011000100012>
- Santos, A. J., Sarmento, H. B., Lima, C. P., & Prada, A. R. R. (2023). Tradução e validação do questionário de clima motivacional percebido no desporto 2 (PMCSQ-2) para português. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 70(4), 45-56.
- Schunk, D. H. (1995). Self-efficacy and education and instruction. *Self-efficacy, adaptation, and adjustment: Theory, research, and application*, 281-303.
- Silva, L. F. F.; Baracat, E. C. (2016) *Medical education - historic perspective*. *Ver. Med. São Paulo*, 95, 28-36.
<http://dx.doi.org/10.11606/issn.1679-9836.v95ispe1p28-36>
- Tawfeek, E. E., Gouda, E., Elaraby, S., & Talaat, W. (2021). Relationship between the role of the

clinical instructor, students' self-efficacy, and academic achievement. *Suez Canal University Medical Journal*, 24(2), 164-177.

<https://www.academia.edu/68315771>.

Turan, S., et al. (2013). Studying self-efficacy beliefs in medical education. *Procedia - Social and Behavioral Sciences*, 93, 1311-1314.

<https://doi.org/10.1016/j.sbspro.2013.10.034>

Uriarte, J. J., Iglesias, N., Penas, P., Moreno-Calvete, M. C., Álvarez-González, A., & Iraurgi, I. (2024). Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 2(72), 145-156.

Vitoria, F., Almeida, L. S., & Primi, R. (2006). Unidimensionalidade em testes psicológicos: conceito, estratégias e dificuldades na sua avaliação. *Psic: revista da Vetor Editora*, 7(1), 01-07.

<http://pepsic.bvsalud.org/pdf/psic/v7n1/v7n1a02.pdf>

Anexo

Escala MASS (Versão Brasileira)

Prezado (a) participante:

O Sr. (a) está sendo convidado (a) a participar desta pesquisa que tem como objetivo adaptar a Escala de Autoeficácia de Realização Médica (MASS – *MEDICAL ACHIEVEMENT SELF-EFFICACY SCALE*) para o contexto universitário em Medicina.

Por gentileza, indique sua opinião sobre cada um dos itens abaixo, respondendo a pergunta: “Sou capaz de...”. Marque sua resposta em uma escala de 1 a 5, representando um contínuo entre pouco e muito. É relevante mencionar que não há respostas certas ou erradas. Suas respostas são confidenciais.

“Sou capaz de...”

1. Aplicar as habilidades que aprendi até agora em um paciente.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
2. Ter conhecimento suficiente sobre os fatores sociais que influenciam os problemas de saúde dos pacientes.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
3. Pesquisar a literatura relevante para um problema de saúde por meio eletrônico.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
4. Aplicar adequadamente as etapas subsequentes de diagnóstico e tratamento a um problema clínico.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
5. Estabelecer uma comunicação adequada com um paciente em uma situação de conflito.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
6. Ter um bom conhecimento dos aspectos médicos que tem sido discutido pelas ciências humanas (por exemplo, filosofia, arte, ...).
() 1 Pouco () 2 () 3 () 4 () 5 Muito
7. Ter conhecimento adequado das ciências médicas básicas.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
8. Fazer uma escolha com custo-benefício adequado ao usar meios técnicos para diagnóstico ou tratamento.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
9. Escrever um artigo científico confiável sobre um assunto relacionado à saúde.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
10. Escolher/elaborar um projeto científico para solucionar um problema de pesquisa médica.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
11. Assumir uma posição pessoal fundamentada sobre os aspectos éticos quando um paciente solicita a ortotanásia (morte natural, sem antecipação ou prolongamento).
() 1 Pouco () 2 () 3 () 4 () 5 Muito
12. Colaborar em condições de igualdade com colegas de outras áreas da saúde.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
13. Lidar com meus sentimentos de ansiedade quando eles aparecem em certas situações clínicas.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
14. Lidar com um problema de saúde na sociedade preventivamente.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
15. Estruturar as informações que obtenho de um paciente em uma consulta.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
16. Analisar um problema de saúde de um paciente em um grupo.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
17. Reconhecer sinais e sintomas de *burnout* (estresse relacionado ao trabalho) no meu ato profissional.
() 1 Pouco () 2 () 3 () 4 () 5 Muito
18. Lidar com incidentes críticos (eventos inesperados e estressantes) enquanto presto cuidados de saúde.
() 1 Pouco () 2 () 3 () 4 () 5 Muito