

**EVIDÊNCIAS DE VALIDADE DE CONTEÚDO DA ESCALA DE ATITUDES  
PERANTE AS AVALIAÇÕES EXTERNAS APLICADAS EM LARGA ESCALA  
(EAAE)**

***EVIDENCIA DE LA VALIDEZ DEL CONTENIDO DE LA ESCALA DE ACTITUD  
HACIA LAS EVALUACIONES EXTERNAS A GRAN ESCALA (EAAE)***

***EVIDENCE OF CONTENT VALIDITY OF THE SCALE OF ATTITUDES TOWARDS  
EXTERNAL ASSESSMENTS APPLIED ON A LARGE SCALE (EAAE)***

Denilson Junio Marques SOARES<sup>1</sup>

Talita Emidio Andrade SOARES<sup>2</sup>

Wagner dos SANTOS<sup>3</sup>

**RESUMO:** Este estudo teve como objetivo desenvolver e verificar as evidências de validade de conteúdo da Escala de Atitudes perante as Avaliações Externas aplicadas em larga escala (EAAE). Esse instrumento foi construído de modo a capturar o que os professores da educação básica pensam, sentem e como se comportam mediante esse tipo de avaliação. Trata-se de um estudo com abordagem quantitativa, estruturado em duas vertentes: avaliação de juízes especialistas com base no cálculo do Coeficiente de Validade de Conteúdo (CVC); e aplicação piloto, realizada com uma amostra da população-alvo. A versão final da escala contou com 30 itens, cujo CVC geral mostrou-se em 0,92, evidenciando sua validade de conteúdo e sinalizando para a clareza de linguagem, pertinência prática e relevância teórica do instrumento. Espera-se que este trabalho abra perspectivas para novas pesquisas, com vistas a dar continuidade ao processo de validação da EAAE e, conseqüentemente, possibilitar seu uso em estudos sobre avaliação educacional.

**PALAVRAS-CHAVE:** Escala de atitudes. Avaliações externas aplicadas em larga escala. Validade de conteúdo. Avaliação educacional. Psicometria.

**RESUMEN:** Este artículo presenta el proceso de construcción y evidencia de validez de contenido de la Escala de Actitudes hacia las Evaluaciones Externas aplicadas a gran escala (EAAE). Este instrumento fue desarrollado con el fin de captar lo que piensan, sienten y se comportan los docentes de educación básica a través de este tipo de evaluación. Se trata de un estudio con enfoque cuantitativo, estructurado en dos maneras: evaluación por jueces expertos, mediante el cálculo del Coeficiente de Validez de Contenido (CVC); y aplicación previa, realizada con una muestra de la población objeto de estudio. La versión final de la escala contó con 30 ítems, cuyo CVC global fue de 0,92, evidenciando su validez de contenido y señalando

<sup>1</sup> Instituto Federal de Minas Gerais (IFMG), Piumhi – MG – Brasil. Professor. Doutorando em Educação (UFES). ORCID: <https://orcid.org/0000-0003-3075-3532>. E-mail: [denilson.marques@ifmg.edu.br](mailto:denilson.marques@ifmg.edu.br)

<sup>2</sup> Universidade Federal do Espírito Santo (UFES), Vitória – ES – Brasil. Doutoranda em Educação. ORCID: <https://orcid.org/0000-0003-2692-4941>. E-mail: [talitaeandrade@gmail.com](mailto:talitaeandrade@gmail.com)

<sup>3</sup> Universidade Federal do Espírito Santo (UFES), Vitória – ES – Brasil. Professor no Programa de Pós-Graduação em Educação. Líder do Instituto de Pesquisa em Educação e Educação Física (Proteoria/UFES). Doutorado em Educação (UFES). ORCID: <https://orcid.org/0000-0002-9216-7291>. E-mail: [wagnercefd@gmail.com](mailto:wagnercefd@gmail.com)

*la claridad del lenguaje, la relevancia práctica y teórica del instrumento. Se espera que este trabajo aporte datos importantes para futuras investigaciones, con miras a continuar el proceso de validación de la EAAE, y en consecuencia, posibilitar su uso en estudios acerca de la evaluación educativa.*

**PALABRAS CLAVE:** *Escala de actitudes. Evaluaciones externas a gran escala. Validez de contenido. Evaluación educativa. Psicometría.*

**ABSTRACT:** *This study presents the process of development of the Scale of Attitudes towards External Assessments applied on a large scale (EAAE) and evidence of its content validity. This instrument was developed in order to capture what basic education teachers think, feel and how they behave through this type of assessment. This is a study with a quantitative approach, structured in two stages: evaluation by expert, using the Content Validity Index (CVI); and pilot application was conducted with a sample of the target population. The final version of the EAAE was reduced to 30 items. The overall CVI of the scale was 0.92, evidencing its content validity and signaling the clarity of language, practice relevance and theoretical relevance of the instrument. It is hoped that this study will open perspectives for further research, with a view to continuing the EAAE validation process and, consequently, enabling its use in studies on educational assessment.*

**KEYWORDS:** *Attitude scales. External assessments applied on a large scale. Content validity. Educational assessment. Psychometry.*

## **Introdução**

No âmbito educacional, a avaliação se configura em variadas modalidades: avaliação de aprendizagens; de currículos e programas; de sistemas educativos; de profissionais; e de políticas públicas (AFONSO, 2000). Conforme Stadler (2017), em todas, a avaliação tem o papel crucial de analisar a qualidade da educação por meio do acompanhamento do processo educativo ou da leitura de seus resultados.

Quando esse processo se realiza por agentes externos à escola, a avaliação é dita externa. Na maioria das vezes, sua aplicação ocorre em larga escala, ou seja, para um grande número de pessoas. Em síntese, as avaliações externas aplicadas em larga escala têm sido utilizadas pelos órgãos governamentais como ferramentas balizadoras de políticas, projetos, programas e ações educativas escolares e comunitárias, podendo ser classificadas conforme: sua natureza, referindo-se à sua abrangência; seus propósitos/finalidades; e seus efeitos (BAUER; ALAVARSE; OLIVEIRA, 2015; ORFIELD; WALD, 2000).

A classificação quanto aos efeitos considera os impactos que a avaliação traz para a sua população-alvo. Nessa vertente, as avaliações de alto impacto (*high stakes tests*) podem oferecer subsídios para os processos de tomada de decisões e, assim, fazer ressurgir

consequências conspícuas para os atores do processo escolar. Em contrapartida, as avaliações de baixo impacto (*low stakes tests*) não oferecem efeitos direta ou indiretamente para esses sujeitos (BAUER; ALAVARSE; OLIVEIRA, 2015; ORFIELD; WALD, 2000).

No Brasil, a Política Nacional de Avaliação da Educação Básica, instituída por meio da publicação da Portaria nº 458, de 5 de maio de 2020 (BRASIL, 2020), é composta por três avaliações externas aplicadas em larga escala: o Sistema de Avaliação da Educação Básica (Saeb), criado em 1990; o Exame Nacional do Ensino Médio (Enem), criado em 1998; e o Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), criado em 2002. O art. 8º e o art. 18 da referida portaria conferem ao Saeb e ao Enem, respectivamente, o mesmo objetivo geral: “[...] aferir o domínio das competências e das habilidades esperadas ao longo da educação básica, de acordo com a Base Nacional Comum Curricular - BNCC e as correspondentes diretrizes curriculares nacionais” (BRASIL, 2020, p. 2). Ao Encceja, conforme o art. 12, cabe aferir “[...] as competências e as habilidades exigidas para a conclusão do ensino fundamental e médio, conforme o caso” (BRASIL, 2020, p. 2).

Em comum, todas elas possuem o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), autarquia federal vinculada ao Ministério da Educação do Brasil (MEC), como agente externo responsável por sua organização e gerenciamento. Ademais, essas avaliações, de natureza nacional, podem ser consideradas de alto impacto para os estudantes a quem se destinam e, conseqüentemente, aos demais sujeitos que constituem o cotidiano das escolas. O Encceja, por exemplo, pode ser utilizado com o propósito de certificar diferentes níveis de ensino; e o Saeb e o Enem, como mecanismo único, alternativo ou complementar para o acesso desses estudantes à educação superior.

Também é preciso ressaltar que, a partir dos resultados obtidos nessas avaliações, é possível realizar um diagnóstico da educação básica brasileira e de fatores que podem interferir no desempenho do estudante, pois, juntamente com elas, são aplicados questionários que permitem explicar os níveis de aprendizagem demonstrados pelos estudantes avaliados a partir das diversas informações contextuais coletadas. Com isso, o uso político desses instrumentos tem ocorrido há décadas e é frequente a apropriação dos seus resultados para delinear as políticas educacionais contemporâneas (MILITÃO; SANTANA; PERBONI, 2019).

Contudo, ainda há uma lacuna no que diz respeito aos estudos voltados a analisar as atitudes dos professores perante essas avaliações. Em um levantamento realizado em janeiro de 2022, no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), mediante consulta expandida às bases de dados eletrônicos *Scientific*

*Electronic Library Online (SciELO), Scopus e Web of Science (WoS)*, utilizando os descritores [“Avaliação externa” OR “Avaliações externas”] AND [“atitude de professores” OR “atitude de docentes”], não foram mapeadas pesquisas produzidas com essa finalidade.

Desse modo, este artigo tem o objetivo de apresentar a Escala de Atitudes perante as Avaliações Externas aplicadas em larga escala (EAAE) e as análises realizadas a fim de encontrar evidências de validade do seu conteúdo. Essa é uma das etapas iniciais do seu processo de construção.

O construto atitudes, utilizado nesse processo, tem origens no início do século XX, a partir do estudo dos sociólogos Thomas e Znaniecki (1918), que o conceituam como um estado individual e determinante para a ação dos indivíduos. De acordo com Sarti, Vendramini e Camilo (2021), para explicá-lo, há diferentes modelos na literatura. O tripartido é o mais relevante, sobretudo por abarcar múltiplos fatores psicológicos. Esse modelo considera as seguintes dimensões inter-relacionadas: cognitiva, afetiva e comportamental.

Segundo Camilo (2021), a dimensão cognitiva abrange as percepções, pensamentos e conceitos em relação ao objeto investigado, buscando captar o que os sujeitos sabem (ou acreditam que sabem) sobre ele; a dimensão afetiva, por sua vez, refere-se às emoções e sentimentos desses sujeitos; e a comportamental abarca suas ações e/ou intenções de agir diante do objeto. Nessa vertente, a EAAE foi desenvolvida de modo a capturar o que os(as) professores(as) da educação básica, que se constituem como a população-alvo da escala, pensam (dimensão cognitiva), sentem (dimensão afetiva) e como se comportam (dimensão comportamental) mediante as avaliações externas aplicadas em larga escala no Brasil.

Dessa forma, o presente artigo encontra-se estruturado em três seções, para além desta introdução. Na primeira, apresenta-se o método utilizado para a realização da pesquisa. Em seguida, são evidenciados os resultados alcançados. Posteriormente, encerrando o artigo, é realizada uma discussão acerca dos resultados obtidos, acompanhada das considerações finais, seguidas das referências.

## Método

### Participantes

Os itens da EAAE foram inicialmente submetidos à avaliação de dez juízes especialistas: oito professores doutores e dois doutorandos, todos vinculados a Programas de Pós-Graduação em Educação e/ou Psicologia das cinco regiões brasileiras. A seleção desses profissionais se deu intencionalmente, considerando a reconhecida competência na área de avaliação educacional e/ou psicológica.

A partir dos resultados obtidos na análise dos juízes especialistas, os itens pré-selecionados para compor a escala passaram por uma avaliação semântica, por meio da aplicação do instrumento para uma amostra de 30 professores da educação básica (aplicação piloto). O propósito dessa etapa foi verificar se os itens eram compreensíveis quanto à linguagem, considerando os sujeitos pertencentes à população-alvo da escala, conforme indicado por Glaser e Strauss (2017). O tamanho da amostra se deu pelo critério da saturação (GLASER; STRAUSS, 2017).

### Instrumento

A EAAE (em processo de elaboração) contou, inicialmente, com 42 itens elaborados por pesquisadores do Instituto de Pesquisa em Educação e Educação Física da Universidade Federal do Espírito Santo, a partir de estudos que consideram a definição constitutiva e operacional do fenômeno investigado (PASQUALI, 2010), desenvolvidos desde 2017 no âmbito do instituto.

Para tanto, foi considerado o construto atitudes, composto pelas dimensões cognitiva (15 itens), afetiva (12 itens) e comportamental (15 itens), no intuito de capturar o que os(as) professores(as) da educação básica (público-alvo) pensam, sentem e como se comportam, respectivamente, em relação às avaliações externas que são aplicadas em larga escala no Brasil.

Cada dimensão é acompanhada por uma frase norteadora: para a dimensão cognitiva é solicitada uma resposta referente ao que se acredita (crenças, conhecimentos, informações e/ou opiniões) em relação às avaliações externas aplicadas em larga escala; para a dimensão afetiva, é solicitada uma resposta baseada nos sentimentos; e, para a dimensão comportamental, a resposta deve estar alicerçada nas ações cotidianas em relação ao objeto.

Os itens que compõem o instrumento foram elaborados na forma de asserções e, como opção de resposta, eles se estruturam em uma escala do tipo *Likert* de cinco pontos de

concordância, variando de 1 (discordo fortemente) a 5 (concordo fortemente). Em síntese, é esperado que menores (maiores) pontuações indiquem atitudes menos (mais) positivas/favoráveis em relação ao objeto investigado.

## **Procedimentos**

As análises realizadas foram viabilizadas mediante envio, por correio eletrônico, de uma carta-convite, acompanhada da solicitação de avaliação dos itens previamente elaborados. Dessa forma, após a aceitação e anuência do participante ao Termo de Consentimento Livre e Esclarecido (TCLE), disponibilizou-se o acesso ao formulário de análise. O sigilo dos participantes foi mantido durante todo o processo.

Na avaliação, cada juiz especialista deveria atribuir um valor ordinal, entre 1 e 5, de acordo com a qualidade do item. Foram considerados três aspectos: a clareza de linguagem (1. nada claro – 5. totalmente claro); a pertinência prática (1. nada pertinente – 5. totalmente pertinente); e a relevância teórica (1. nada relevante – 5. totalmente relevante) dos itens. Esse procedimento assumiu como referência os estudos de Hernandez-Nieto (2002), Cassepp-Borges, Balbinotti e Teodoro (2010) e Pacico (2015).

Conforme indicado por Cassepp-Borges, Balbinotti e Teodoro (2010), a clareza de linguagem consiste na análise da linguagem utilizada nos itens, considerando as características do público-alvo. A pertinência prática, por sua vez, objetiva analisar se o item é, de fato, importante na constituição do instrumento. Já a relevância teórica busca avaliar o grau de associação entre o item e a teoria que lhe oferece suporte.

Para a análise semântica, foi elaborado um questionário em que os participantes da aplicação piloto puderam classificar a linguagem adotada como clara, difícil ou incompreensível. Nessa etapa, assumiu-se como subsídio teórico o estudo de Pasquali (2010).

## **Análise estatística**

A análise quantitativa foi realizada calculando-se a média aritmética e o desvio padrão das respostas atribuídas pelos juízes especialistas para cada item, em cada aspecto analisado. Em síntese, quanto maior (menor) o valor para a primeira, mais (menos) bem avaliado é, em média, o item. Já em relação ao desvio padrão, valores altos (baixos) indicam uma menor (maior) regularidade de respostas, sinalizando para uma possível divergência de opiniões entre os avaliadores.

Para determinar o nível de concordância entre os juízes especialistas, foi utilizado o Coeficiente de Validade de Conteúdo (CVC), proposto por Hernández-Nieto (2002). Trata-se de um indicador utilizado para calcular o nível de consenso, mediante o cálculo da média do valor atribuído pelos juízes em cada item do instrumento, expresso em termos percentuais, subtraindo-se o viés de resposta. Em termos matemáticos:

$$CVC = \frac{\sum_{i=1}^K \frac{x_i}{x_M}}{K} - \left(\frac{1}{K}\right)^K,$$

em que  $x_i$  representa o valor  $x$  atribuído ao item pelo juiz  $i; K$  representa o número de juízes; e  $x_M$ , a pontuação máxima da escala utilizada.

Além do CVC para clareza de linguagem, pertinência prática e relevância teórica, a média para todos os itens e o CVC global da escala também foram calculados. Conforme indicado por Polit e Beck (2006), foram considerados válidos itens com CVC igual ou superior a 80%. Todas as análises foram conduzidas com o auxílio do *software* Microsoft Excel®.

## Resultados

A partir das avaliações dos juízes especialistas, dos 42 itens previamente elaborados, 12 foram excluídos por apresentarem informações repetidas ou CVC inferior a 80%. Dos demais, cinco foram reformulados considerando as sugestões dos avaliadores a fim de dar objetividade à escrita ou facilitar a sua compreensão.

Dessa forma, 30 itens foram previamente selecionados para compor a EAAE: 12 do componente cognitivo, oito do componente afetivo e dez do componente comportamental. A análise semântica, realizada a partir da aplicação piloto do instrumento, não indicou problemas quanto à sua compreensão para a amostra do público-alvo investigada (N=30).

O escore total obtido pelos respondentes pode variar de 30 a 150 pontos, com ponto médio igual a 90 pontos. Assim, pontuações acima (abaixo) desse valor indicam atitudes mais (menos) positivas do que negativas dos professores em relação às avaliações externas aplicadas em larga escala.

Os Quadros 1, 2 e 3 trazem esses itens, organizados por dimensão e acompanhados da pergunta norteadora, média e desvio padrão dos valores da escala *Likert* atribuídos pelos juízes,

bem como o CVC para cada critério considerado (clareza de linguagem, pertinência prática e relevância teórica do item).

**Quadro 1** – Média, desvio padrão e CVC dos itens do componente cognitivo da EAAE

	Componente Cognitivo Itens	C. L.		P. P		R. T.	
		M (DP)	CVC	M (DP)	CVC	M (DP)	CVC
Resposta baseando-se no que você acredita (seus conhecimentos, opiniões e/ou crenças) em relação às avaliações externas aplicadas em larga escala.	Avaliam adequadamente a qualidade do ensino e da aprendizagem	4,50 (0,85)	0,90	4,60 (0,52)	0,92	4,90 (0,32)	0,98
	Constituem o principal instrumento de monitoramento da efetividade de políticas educacionais	4,70 (0,67)	0,94	4,80 (0,63)	0,96	4,90 (0,32)	0,98
	Cumprem satisfatoriamente o propósito de medir os níveis de aprendizagem dos alunos	4,20 (1,03)	0,84	4,40 (1,35)	0,88	4,40 (1,35)	0,88
	São úteis para selecionar adequadamente estudantes para outras etapas/níveis de ensino	4,70 (0,67)	0,94	4,50 (1,27)	0,90	4,60 (1,26)	0,92
	Têm impacto na seleção dos conteúdos ensinados na escola	4,40 (0,84)	0,88	4,30 (1,34)	0,86	4,10 (1,37)	0,82
	Medem adequadamente o nível de conhecimento dos estudantes em relação aos conteúdos programáticos previstos na Base Nacional Comum Curricular (BNCC)	4,40 (0,84)	0,88	4,90 (0,32)	0,98	5,00 (0,00)	1,00
	Suas matrizes curriculares conferem igual valor a todos os conhecimentos previstos na Base Nacional Comum Curricular (BNCC)	4,60 (1,26)	0,92	4,20 (1,69)	0,84	4,20 (1,69)	0,84
	Permitem a análise dos aspectos sociais e culturais dos alunos	4,20 (1,03)	0,84	4,80 (0,63)	0,96	4,80 (0,63)	0,96
	São instrumentos necessários para reduzir as desigualdades educacionais	4,80 (0,63)	0,96	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	São importantes instrumentos de prestação de contas à sociedade	4,80 (0,63)	0,96	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	São importantes instrumentos de monitoramento da prática docente	4,90 (0,32)	0,98	4,40 (1,35)	0,88	4,40 (1,35)	0,88
	São instrumentos eficientes para promover bonificações financeiras aos profissionais da educação	4,50 (0,71)	0,90	4,40 (0,84)	0,88	4,50 (0,85)	0,90
	Total:	4,56 (0,79)	0,91	4,61 (0,86)	0,92	4,65 (0,76)	0,93

Nota: M = média aritmética; D. P. = desvio padrão; C. L. = clareza de linguagem; P. P. = pertinência prática; R. T. = relevância teórica.

Fonte: Elaborado pelos autores (2021)

**Quadro 2** – Média, desvio padrão e CVC dos itens do componente afetivo da EAAE

	Componente Afetivo	C. L.		P. P		R. T.	
		M (DP)	CVC	M (DP)	CVC	M (DP)	CVC
Resposta baseando-se em seus sentimentos em relação às avaliações externas aplicadas em larga escala	Tenho apreço por esse tipo de avaliação	5,00 (0,00)	1,00	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Anseio por conhecer os seus resultados	4,80 (0,42)	0,96	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Confio em seus resultados	5,00 (0,00)	1,00	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Confio nas comparações entre escolas/estudantes que são feitas a partir delas	4,60 (0,84)	0,92	4,70 (0,67)	0,94	4,80 (0,63)	0,96
	Confio nas certificações de conclusão de um nível/etapa de ensino que são realizadas a partir delas	4,70 (0,67)	0,94	4,80 (0,63)	0,96	4,80 (0,63)	0,96
	Sinto que meus conhecimentos são valorizados por elas	5,00 (0,00)	1,00	4,90 (0,32)	0,98	5,00 (0,00)	1,00
	Percebo que a disciplina que leciono é valorizada por elas	4,40 (0,84)	0,88	4,30 (0,82)	0,86	4,50 (0,53)	0,90
	Gosto de orientar meu trabalho a partir dos seus resultados	4,90 (0,32)	0,98	4,40 (1,35)	0,88	4,40 (1,35)	0,88
	Total:	4,77 (0,44)	0,95	4,73 (0,54)	0,95	4,81 (0,39)	0,96

Nota: M = média aritmética; D. P. = desvio padrão; C. L. = clareza de linguagem; P. P.= pertinência prática; R. T. = relevância teórica.

Fonte: Elaborado pelos autores (2021)

**Quadro 3** – Média, desvio padrão e CVC dos itens do componente comportamental da EAAE

	Componente Comportamental	C. L.		P. P		R. T.	
		M (DP)	CVC	M (DP)	CVC	M (DP)	CVC
Resposta baseando-se em suas ações cotidianas em relação às avaliações externas aplicadas em larga escala	Planejo minhas aulas a partir dos seus resultados	5,00 (0,00)	1,00	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Incorporo seus resultados em minhas ações cotidianas na sala de aula	4,50 (0,85)	0,90	4,60 (1,26)	0,92	4,60 (1,26)	0,92
	Organizo o cronograma escolar para valorizar sua realização	5,00 (0,00)	1,00	4,60 (1,26)	0,92	4,60 (1,26)	0,92
	Priorizo o ensino dos conteúdos que são mais frequentes nessas avaliações	4,70 (0,67)	0,94	4,90 (0,32)	0,98	4,90 (0,32)	0,98
	Converso com estudantes sobre sua importância	5,00 (0,00)	1,00	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Recomendo a participação nessas avaliações	4,50 (1,08)	0,90	4,60 (0,84)	0,92	4,60 (0,84)	0,92
	Apresento um retorno ( <i>feedback</i> ) sobre os resultados obtidos em sala de aula	4,90 (0,32)	0,98	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Resolvo questões de avaliações anteriores em sala de aula	4,80 (0,42)	0,96	4,90 (0,32)	0,98	4,80 (0,63)	0,96
	Utilizo questões anteriores em minhas avaliações escolares (internas)	4,70 (0,67)	0,94	5,00 (0,00)	1,00	5,00 (0,00)	1,00
	Participo de formação específica para uma melhor utilização dos seus resultados	4,00 (1,49)	0,80	4,10 (1,37)	0,82	4,20 (1,40)	0,84

Total:	4,68 (0,61)	0,94	4,74 (0,60)	0,95	4,77 (0,57)	0,950
--------	----------------	------	----------------	------	----------------	-------

Nota: M = média aritmética; D. P. = desvio padrão; C. L. = clareza de linguagem; P. P.= pertinência prática; R. T. = relevância teórica.

Fonte: Elaborado pelos autores (2021)

Além do CVC para cada critério, o CVC da escala como um conjunto uniforme também foi calculado. Assim, para o aspecto clareza de linguagem, o CVC geral foi de 0,92, para a pertinência prática foi de 0,93 e para a relevância teórica 0,94. O CVC global da escala foi de 0,93, conferindo sua validade de conteúdo conforme os métodos de análise utilizados nesta pesquisa.

Quanto às análises descritivas, percebe-se uma melhor média, na avaliação dos juízes, para o critério relevância teórica (M=4,73), em comparação com os demais (M=4,70 em ambos). Quanto ao desvio padrão, essa avaliação é mais heterogênea para o critério pertinência prática (D. P. = 0,64) e mais homogênea para a clareza de linguagem (D. P. = 0,57). Para o critério relevância teórica, o desvio padrão calculado foi 0,60.

## Discussão e considerações finais

Este artigo foi desenvolvido com o intuito de apresentar o processo de construção e evidências de validade de conteúdo da Escala de Atitudes perante as Avaliações Externas aplicadas em larga escala (EAAE). Esse instrumento foi construído a partir da adoção do modelo tripartido do construto atitudes, de modo a capturar o que os professores da educação básica pensam, sentem e como se comportam mediante as avaliações externas aplicadas em larga escala.

O processo de construção e a busca por evidências de validação de conteúdo da EAAE foram pautados rigorosamente em procedimentos reconhecidos pela literatura especializada, consultada para o desenvolvimento do instrumento. A utilização do CVC, na análise dos juízes especialistas, e a análise semântica demonstraram bons resultados, fornecendo subsídios para a conclusão de que o instrumento apresenta evidências de validade de conteúdo, sinalizando para a clareza de linguagem, pertinência teórica e relevância prática dos itens que o compõem.

Nessa vertente, é preciso ressaltar que quatro itens apresentaram concordância de 100% em todos os aspectos avaliados: dois do componente afetivo (“Tenho apreço por esse tipo de avaliação” e “Confio em seus resultados”) e dois do comportamental (“Planejo minhas aulas a partir dos seus resultados” e “Converso com estudantes sobre sua importância”), indicando uma avaliação perfeita unânime entre os avaliadores. É esperado, assim, teoricamente, que esses

itens tragam maiores informações a respeito das atitudes dos professores da educação básica em relação às avaliações externas aplicadas em larga escala.

Entretanto, para confirmar essa hipótese, bem como para avaliar outros aspectos psicométricos da escala, é preciso que novas análises de validade sejam realizadas, como a baseada na estrutura interna do instrumento, nas relações com medidas externas ou no padrão de resposta aos itens. Isso se fará em estudos posteriores, considerando que a EAAE ainda se encontra em construção.

Como principal limitação deste estudo, ressalta-se o fato de as análises terem sido realizadas no formato on-line, o que dificulta o controle de variáveis, como cansaço e distração dos avaliadores, que podem comprometer os resultados encontrados. Por outro lado, foi possível contar com a colaboração de profissionais que atuam em todas as regiões do país, proporcionando uma avaliação consideravelmente abrangente dos itens da escala.

Por fim, ressalta-se que, quando finalizada, a EAAE poderá subsidiar uma nova agenda de pesquisa no âmbito educacional, contribuindo para o crescimento e aprimoramento da produção científica. A partir de sua aplicação para o público-alvo, algumas questões poderão ser respondidas, por exemplo: a) Como os professores da educação básica se posicionam em relação às avaliações externas aplicadas em larga escala? b) Há relação significativa entre variáveis sociais/demográficas/econômicas e atitudes mais positivas quanto a elas? c) Atitudes mais positivas influenciam os resultados educacionais alcançados? Entre outras.

Com isso, a EAAE poderá oferecer subsídios para o estabelecimento de soluções eficientes, estratégicas e inovadoras voltadas para a melhoria da qualidade da educação, evidenciada por meio das referidas avaliações. Dessa forma, ressalta-se que a construção dessa escala é de extrema relevância, sobretudo diante da escassez de estudos que abordam a temática investigada.

**AGRADECIMENTOS:** À Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (Fapes) pelo apoio ao desenvolvimento desta pesquisa.

## REFERÊNCIAS

- AFONSO, A. J. **Avaliação educacional**: Regulação e emancipação para uma sociologia das políticas avaliativas contemporâneas. 2. ed. São Paulo: Cortez, 2000.
- BAUER, A.; ALAVARSE, O. M.; OLIVEIRA, R. P. Avaliação em larga escala: Uma sistematização do debate. **Educação e Pesquisa**, São Paulo, v. 41, n. esp., p. 1367-1382, dez. 2015. Disponível em: <http://www.scielo.br/pdf/ep/v41nspe/1517-9702-ep-41-spe-1367.pdf>. Acesso em: 03 jan. 2022.
- BRASIL. **Portaria n. 458, de 5 de maio de 2020**. Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica. Brasília, DF: Ministro de Estado da Educação, 2020. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-458-de-5-de-maio-de-2020-255378342>. Acesso em: 05 jan. 2022.
- CAMILO, C. C. **Construção e estudos psicométricos da Escala de Atitude frente à atuação do psicólogo**. 2021. Tese (Doutorado em Psicologia) – Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Campinas, 2021. Disponível em: <https://www.usf.edu.br/galeria/getImage/427/4000894494237433.pdf>. Acesso em: 5 jan. 2022.
- CASSEPP-BORGES, V.; BALBINOTTI, M. A. A.; TEODORO, M. L. M. Tradução e validação de conteúdo: Uma proposta para a adaptação de instrumentos. *In*: PASQUALI, L. **Instrumentação psicológica**: Fundamentos e práticas. Porto Alegre: Artmed, 2010.
- GLASER, B. G.; STRAUSS, A. L. **Discovery of grounded theory**: Strategies for qualitative research. New York: Aldine de Gruyter, 2017.
- HERNÁNDEZ-NIETO, R. A. **Contribuciones al análisis estadístico**. Mérida, Venezuela: Universidad de Los Andes/iesinfo, 2002.
- MILITÃO, A. N.; SANTANA, A. C. M.; PERBONI, F. A expansão dos sistemas avaliativos e a teoria crítica: Um diálogo (im)pertinente. **Revista Ibero-Americana de Estudos em Educação**, Araraquara, v. 14, n. esp. 4, p. 2046-2056, 2019. Disponível em: <https://periodicos.fclar.unesp.br/iberoamericana/article/view/12927>. Acesso em: 22 dez. 2021.
- ORFIELD, G.; WALD, J. Testing, testing: The high-stakes testing mania hurts poor and minority students the most. **The Nation**, New York, v. 270, n. 22, p. 38-40, jun. 2000. Disponível em: <https://thenation.s3.amazonaws.com/pdf/testingtesting2000.pdf>. Acesso em: 05 jan. 2022.
- PACICO, J. C. Como é feito um teste? Produção de itens. *In*: HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. **Psicometria**. Porto Alegre: Artmed, 2015.
- PASQUALI, L. **Instrumentação psicológica**. Brasília, DF: Editora Vetor, 2010.
- POLIT, D. F.; BECK, C. T. The content validity index: Are you sure you know what's being reported? Critique and recommendations. **Research in Nursing & Health**, v. 29, n. 5, p. 489-497, 2006. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nur.20147>. Acesso em: 15 jan. 2022.

SARTI, A. G.; VENDRAMINI, C. M. M.; CAMILO, C. C. Evidências de validade da escala de atitudes de estudantes do ensino superior frente à estatística-EAEst. **Ciência & Educação**, Bauru, v. 27, e21042, 2021. Disponível em: <https://www.scielo.br/j/ciedu/a/zJ9mkwgpzyP5zRZjhsYsfh/abstract/?lang=pt>. Acesso em: 10 dez. 2021.

STADLER, J. C. **Prova Brasil de Matemática do 5º ano do ensino fundamental: Resultados nas plataformas Devolutivas Pedagógicas e QEdU**. Ponta Grossa, 2017. 165 f. Dissertação (Mestrado em Educação) – Universidade Estadual de Ponta Grossa, Ponta Grossa, 2017. Disponível em: <http://tede2.uepg.br/jspui/handle/prefix/2394>. Acesso em: 15 jan. 2022.

THOMAS, W. I.; ZNANIECKI, F. **The Polish peasant in Europe and America: Monograph of an immigrant group**. Chicago: University of Chicago Press, 1918.

### Como referenciar este artigo

SOARES, D. J. M.; SOARES, T. E. A.; SANTOS, W. Evidências de validade de conteúdo da Escala de Atitudes perante as Avaliações Externas aplicadas em larga escala (EAAE). **Revista Ibero-Americana de Estudos em Educação**, Araraquara, v. 17, n. 3, p. 1806-1818, jul./set. 2022. e-ISSN: 1982-5587. DOI: <https://doi.org/10.21723/riaee.v17i3.16191>

**Submetido em:** 26/01/2022

**Revisões requeridas em:** 15/04/2022

**Aprovado em:** 06/05/2022

**Publicado em:** 01/07/2022

**Processamento e editoração:** Editora Ibero-Americana de Educação.

Revisão, formatação, normalização e tradução.