

# Matrizes de Referência de Ciências da Natureza e Ciências Humanas: o processo de validação

Wallace Alves Cabral<sup>1</sup>

Departamento de Ciências Naturais -  
UFSJ  
São João del-Rei, MG, Brasil  
wallacecabral@ufsj.edu.br

Miriam Raquel Piazzini  
Machado<sup>2</sup>

Colégio de Aplicação João XXIII - UFJF  
Juiz de Fora, MG, Brasil  
miriam.machado@ufjf.br

## Resumo

O trabalho aqui apresentado, inserido na Fundação CAEd – UFJF, tem como objetivo expor duas pesquisas desenvolvidas por dois professores, voltadas à validação de Matrizes de Referência de Ciências da Natureza e de Ciências Humanas no Ensino Fundamental. As Matrizes da Fundação foram revisadas com base na Base Nacional Comum Curricular (BNCC) e nas referências específicas de cada área. Após essa revisão, os verbos foram adequados à Taxonomia de Bloom revisada. Em seguida, as Matrizes passaram por um processo de validação, inicialmente realizado por um corpo técnico especializado em avaliação e, posteriormente, por professores da Educação Básica de diferentes regiões do Brasil. Esse processo de validação é uma prática ainda recente na realidade educacional brasileira e contribui para a construção de Matrizes mais alinhadas às necessidades dos estudantes e de seus professores.

Palavras-chave: Matrizes de Referência; Ciências da Natureza, Ciências Humanas; Ensino Fundamental.

---

<sup>1</sup> À FUNDAÇÃO CENTRO DE POLÍTICAS PÚBLICAS E AVALIAÇÃO DA EDUCAÇÃO – FUNDAÇÃO CAEd - UFJF, que aprovou e financiou a pesquisa. Ao PPEdu da UFSJ, por meio do PDPG/CAPES, pelo incentivo e apoio financeiro à participação no evento.

<sup>2</sup> Aos professores que se dispuseram voluntariamente a participar da pesquisa, assim como à FUNDAÇÃO CENTRO DE POLÍTICAS PÚBLICAS E AVALIAÇÃO DA EDUCAÇÃO – FUNDAÇÃO CAEd - UFJF, que aprovou e financiou a pesquisa.

## 1 Introdução

A pesquisa apresentada neste trabalho vem sendo desenvolvida desde 2023, quando foram aprovados pela Fundação Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (Fundação CAEd – UFJF) dois projetos relacionados à produção e validação de Matrizes de Referência para as áreas de Ciências da Natureza (Ensino Fundamental II) e Ciências Humanas – História e Geografia (1º Ciclo do Ensino Fundamental I).

A Fundação CAEd tem como objetivo prestar apoio especializado à execução de projetos e serviços relacionados à avaliação educacional em larga escala e ao desenvolvimento de tecnologias de avaliação e gestão da educação pública. Em razão da relevância das Matrizes de Referência para a produção de medidas educacionais, torna-se imprescindível um olhar criterioso sobre elas, por meio de estudos de validade que impactem o trabalho desenvolvido. Nesse sentido, as Matrizes de Referência produzidas e validadas podem nortear os trabalhos desenvolvidos pela Fundação em diferentes frentes de atuação.

Tradicionalmente, no Brasil, as Matrizes de Referência são uma estrutura tabelar com um conjunto de competências e habilidades, ou mesmo uma lista de competência, habilidades e/ou descritores. O processo de validação dessas Matrizes é algo novo na realidade brasileira, e Rêgo e Souza (2022) destacam a importância dos estudos de validade nos relatórios das avaliações educacionais brasileiras, que ainda são incipientes, considerando os impactos que causam na realidade avaliada.

Em se tratando do processo de validação das matrizes aqui tratadas, buscou-se a referência dos Standards, que, segundo Jesus e Rêgo (2022), são referência em todo o mundo em relação aos princípios constituintes da validade. Ainda segundo as autoras, a definição de validade se apresenta com modificações ao longo dos anos:

Atualmente, a validade é definida como “o grau com que a evidência e a teoria apoiam as interpretações dos escores do teste para um propósito específico” (AERA; APA; NCME, 2014; MESSICK, 1989). [...] Sireci (2016) enfatiza que essas definições anteriores permanecem importantes e necessárias; elas não foram anuladas ou substituídas pela teoria moderna; elas foram expandidas (JESUS e RÊGO, 2022, p. 54).

Essa definição de validade ajuda a entender o processo, que não é apenas quantitativo, mas, também, qualitativo. Ademais, esse manual apresenta a evolução dos processos de validade ao longo dos anos e hoje têm-se as evidências de validade e não mais tipos de validade. Hutz (2009), apud Jesus e Rêgo (2022, p. 58), apresenta um quadro sobre as evidências:

**Quadro 1 – As fontes de evidências de validade apresentadas na quarta versão dos Standards e os tipos de dados que são coletados para se levantar cada uma das fontes**

Fonte	Tipos de Dados Coletados
Evidências com base no conteúdo	Acerca da representatividade da matriz e dos itens da medida, investigando se esses consistem em amostras representativas do domínio que se pretende avaliar.
Evidências com base no processo de resposta	Acerca dos processos mentais envolvidos na realização das tarefas propostas pela matriz ou pela teoria.
Evidências com base na estrutura interna	Acerca da representação do construto, com base nas dimensões avaliadas, na qualidade dos itens e na confirmação de hipóteses derivadas da teoria.
Evidências com base nas relações com variáveis externas	Acerca dos padrões de correlação entre os escores da medida e de outras variáveis que medem o mesmo construto ou construtos relacionados (convergência) e variáveis que medem construtos diferentes (divergência). Levantar também dados sobre a capacidade preditiva da medida com relação a outros fatos de interesse direto (critérios externos) que têm importância por si só e associam-se ao propósito direto do uso da medida.
Evidências com base nas consequências da testagem	Acerca das consequências sociais intencionais e não intencionais do uso da medida, com o fim de verificar se sua utilização está surtindo os efeitos desejados na sociedade, de acordo com o propósito para o qual foi criada.

Fonte: HUTZ (2009)

Portanto, busca-se as evidências com base no conteúdo, que dizem sobre a representatividade da Matriz, assim como as evidências com base no processo de resposta, que verifica os processos mentais envolvidos na realização das tarefas.

A partir desse breve contexto, a presente pesquisa, em fase de finalização pelos pesquisadores, tem como objetivo geral expor o processo de validação das Matrizes de Referência produzidas, delimitando os desafios e potencialidades.

## 2 Metodologia

A base metodológica das duas pesquisas aprovadas pela Fundação CAEd foi a mesma: a partir da revisão das Matrizes, elas foram encaminhadas para o processo de validação por juízes. Ou seja, a pesquisa foi desenvolvida em dois grandes eixos, a revisão e a validação das Matrizes de Referência.

No primeiro ano da pesquisa (2023), as Matrizes de Referência de avaliação das áreas de Ciências da Natureza (Ensino Fundamental II) e Ciências Humanas – História e Geografia (1º Ciclo do Ensino Fundamental) da Fundação CAEd foram reformuladas, assimilando a nova Base Nacional Comum Curricular (BNCC) (BRASIL, 2017). Os objetivos das duas pesquisas estavam direcionados ao processo de construção dos descritores e classes, com base na Taxonomia de Bloom revisada (ANDERSON et al., 2001; FERRAZ, BELHOT, 2010). Além disso, na área de Ciências da Natureza, houve a incorporação da perspectiva do Letramento Científico (SANTOS, 2007) na produção das Matrizes, sendo um importante referencial na área de Ensino de Ciências.

Em relação à Ciências Humanas/Geografia, buscou-se referenciar os trabalhos de revisão com base na Geografia Crítica e a importância da Alfabetização Cartográfica, principalmente, de acordo com Callai (2005). Nas Ciências Humanas (História), baseou-se no conceito do sujeito ativo na construção do conhecimento, e tem-se em Oliveira (2005) um dos referenciais.

Já no segundo e terceiro ano da pesquisa (2024 e 2025), iniciou-se o processo de validação, como dito, uma prática ainda recente no contexto das avaliações educacionais no Brasil, conduzido por juízes responsáveis por analisar essas Matrizes. Esse processo está ancorado, principalmente, no manual dos “Standards<sup>3</sup>” (AERA; APA; NCME, 2014).

Para isso, foram construídos formulários no *Google forms* contendo cada uma das Matrizes, e esses formulários foram encaminhados para os especialistas da área, que a partir de uma escala de 1 a 5 (onde 1 é a menor concordância e 5, a maior concordância), deram suas contribuições sobre a relevância, clareza e coerência dos descritores e classes.

Os conceitos foram construídos a partir do trabalho de Jesus (2024), que considera os seguintes critérios: **clareza** – o grau em que o descritor e a classe são facilmente compreensíveis, ou seja, apresentam estrutura gramatical, sintática e semântica adequadas; **coerência** – o quanto a afirmação feita pelo descritor e pela classe representa o currículo da área; e **relevância** – o quanto o descritor e a classe são essenciais para a avaliação do currículo da área.

Primeiramente, as Matrizes foram encaminhadas para aqueles especialistas que se chama de juízes internos, ou seja, aqueles que lidam e produzem as avaliações de larga escala no dia a dia da Fundação CAEd. Após, foram enviados a professores de algumas regiões brasileiras. Esses professores, que se chama juízes externos, participaram voluntariamente da pesquisa.

Após o processo realizado pelos juízes, as matrizes também foram validadas por uma psicometrista da área. Após as validações interna e externa, bem como a análise psicométrica, iniciou-se a análise dos materiais. Os descritores e classes que apresentaram concordância de, no mínimo, 80% entre os juízes foram mantidos, enquanto aqueles com índice inferior a esse valor foram reformulados ou excluídos. Além disso, foi realizada uma adequação dos níveis cognitivos com base na análise psicométrica.

Por fim, feita as análises, as Matrizes passaram por nova revisão, constituindo um material rico para a construção de itens e para a avaliação em larga escala pela Fundação.

### 3 Resultados

Os resultados apontam para uma revisão das Matrizes que dialoga com a BNCC, com a Taxonomia de Bloom revisada, com os especialistas e professores das áreas.

---

<sup>3</sup> Standards for Educational and Psychological Testing (AERA; APA; NCME, 2014). Este livro é uma publicação elaborada por importantes associações americanas, em colaboração com os principais pesquisadores da área de Educação e de psicologia em todo o mundo: a American Educational Research Association (AERA – Associação Americana de Pesquisa Educacional), a American Psychological Association (APA – Associação Americana de Psicologia) e o National Council on Measurement in Educational (NCME – Conselho Nacional sobre Medida Educacional). (JESUS et al, 2022, p. 54).

O processo de validação não foi fácil, porque exigiu um trabalho dos professores além de suas funções já exercidas cotidianamente, e esse trabalho se fez por voluntariado. Essa é uma questão que se colocou como impedimento, pois os formulários foram enviados para muitos professores por *e-mail*, mas muitos não responderam, e considera-se que seria uma tarefa muito árdua para não ser remunerada. Nesse sentido, em próximas edições, recomenda-se o pagamento para todos os juízes envolvidos.

Outra questão que se verifica é a diferença de olhar dos especialistas em avaliação e dos professores. Observa-se que ambos possuíam formação, alguns dos quais doutorado na área. Entretanto, o que poderia se constituir como um desafio para os pesquisadores, tornou-se uma oportunidade de se ter diferentes leituras, e, conseqüentemente, maior possibilidade de adequação e de relevância das matrizes constituídas e reformuladas.

### 3.1 Discussão

De acordo com Jesus (2024, p. 8), “as principais evidências a serem produzidas dizem respeito ao conteúdo e ao processo de resposta”. A validação por base no conteúdo foi feita pelos especialistas, e a validação a respeito do processo de resposta foi feita pela psicometrista.

A evidência baseada no conteúdo pode incluir análises lógicas ou empíricas acerca da representatividade e relevância do conteúdo para a interpretação proposta para os escores do teste. Nesse sentido, a evidência de validade com base no conteúdo também pode ser obtida por meio da análise de juízes. Os juízes avaliam a representatividade da Matriz de Conteúdo, por exemplo, em relação ao propósito do teste (JESUS, 2024, p. 9).

Após esse julgamento, foi possível compreender que alguns descritores e classes precisavam ser excluídos, por não serem tão relevantes para as avaliações, apesar de estarem relacionados com uma importante habilidade da BNCC que devia ser trabalhada em sala de aula, mas difícil de ser mensurada em uma avaliação em larga escala. Outros descritores e classes precisaram ser revisados, por questões relacionadas à clareza e coerência, ou seja, ajustes relacionados aos verbos utilizados, bem como à estrutura gramatical, sintática e semântica.

Um olhar atento também se manifestou em relação ao currículo da área, demonstrando que nem todos os aspectos curriculares tratados na escola são passíveis de serem avaliados em larga escala.

## 4 Conclusões e Considerações Finais

O que se apresentou nesse trabalho quer ser um olhar sobre as matrizes de referência das Ciências da Natureza e das Ciências Humanas, se constituindo uma experiência de validação.

Considera-se que houve uma possibilidade de ouvir aqueles que se tornam os grandes especialistas, os que constroem os instrumentos de avaliação, mas também os professores que estão trabalhando com as disciplinas na escola, que têm a realidade das avaliações em sua prática de sala de aula.

Os trabalhos não são conclusivos, mas apresentam um ensaio rico, demonstrando que as avaliações em larga escala podem e precisam considerar não apenas os estudos na área, mas a experiência constituída no cotidiano.

## 5 Referências

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION – AERA; AMERICAN PSYCHOLOGICAL ASSOCIATION – APA; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION – NCME. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, 2014.

ANDERSON, L.W. (Ed.), KRATHWOHL, D.R. (Ed.), AIRASIAN, P.W., CRUIKSHANK, K.A., MAYER, R.E., PINTRICH, P.R., RATHS, J., WITTROCK, M.C. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman. L., 2001.

BRASIL. Ministério da Educação. *Base Nacional Comum Curricular*. Brasília: MEC, 2017.

CALLAI, H. C. Aprendendo a ler o mundo: a geografia nos anos iniciais do ensino fundamental. *Cadernos CEDES*, v. 25, n. 66, p. 227–247, maio 2005. Disponível em: <<https://www.scielo.br/j/ccedes/a/7mpTx9mbrLG6Dd3FQhFqZYH/?lang=pt>>. Acesso em: 14 ago 2023.

FERRAZ, A. P. C. M. BELHOT, R. V. Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gest. Prod.*, São Carlos, v. 17, n. 2, p. 421-431, 2010.

JESUS, Girlene Ribeiro de. *Proposição de modelo de validação de matrizes de referência: um estudo com a área de língua portuguesa*. Relatório de Pesquisa do CAEd, 2024.

JESUS, Girlene Ribeiro de, RÊGO, Renata Manuely de Lima, SOUZA, Victor Vasconcelos de. A questão da validade na avaliação educacional brasileira. *Ensaio: aval. pol. públ. Educ.*, Rio de Janeiro, v.30, n.114, p. 52-72, jan./mar. 2022.

JESUS, Girlene Ribeiro de. *Proposição de modelo de validação de matrizes de referência: um estudo com a área de língua portuguesa*. *Relatório de Pesquisa da Fundação CAEd*, 2024.

OLIVEIRA, Sandra Regina Ferreira de. O tempo, a criança e o ensino de História. In: ROSSI, Vera Lúcia Sabongi de; ZAMBONI, Ernesta (Org.). *Quanto tempo o tempo tem!* Campinas: Alínea, 2005.

SANTOS, W. L. P. Educação científica na perspectiva de letramento como prática social: funções, princípios e desafios. *Revista Brasileira de Educação*. v. 12, n. 36, 2007.