



Revista
Educar Mais

O coeficiente de validade de conteúdo (CVC) como ferramenta para validação de questionários: exemplo de uso na pesquisa em Educação e Ensino de Ciências

The content validity coefficient (CVC) as a tool for questionnaire validation: an example of use in research in Education and Science Teaching

El coeficiente de validez de contenido (CVC) como herramienta para la validación de cuestionarios: un ejemplo de uso en investigación en Educación y Enseñanza de las Ciencias

Patrick Alves Vizzotto¹ 

RESUMO

Pesquisas em Educação e Ensino de Ciências desempenham um papel fundamental no aprimoramento da qualidade da educação e na compreensão dos desafios enfrentados por professores e alunos. Para coletar dados confiáveis e relevantes é essencial que os instrumentos de pesquisa, como questionários, sejam devidamente validados. O Coeficiente de Validade de Conteúdo (CVC) é uma importante ferramenta para avaliar a validade de conteúdo de um instrumento de coleta de dados, conferindo se os itens que o compõe possuem clareza, pertinência e relevância. O presente trabalho discute a importância do CVC, destacando seus fundamentos e aspectos metodológicos. Por fim, apresenta um exemplo de aplicação prática na validação de um instrumento que visa mensurar o potencial de aulas de Física na promoção de uma aprendizagem significativa.

Palavras-chave: Ensino de Física; Psicometria; Pesquisa quantitativa; Teoria da Aprendizagem Significativa.

ABSTRACT

Research in Education and Science Teaching plays a key role in improving the quality of education and understanding the challenges faced by teachers and students. To collect reliable and relevant data, it is essential that research instruments, such as questionnaires, are properly validated. The Content Validity Coefficient (CVC) is an important tool to assess the content validity of a data collection instrument, checking whether the items that compose it are clear, pertinent, and relevant. The present paper discusses the importance of CVC, highlighting its foundations and methodological aspects. Finally, it presents an example of practical application in the validation of an instrument that aims to measure the potential of Physics classes in promoting meaningful learning.

Keywords: *Physics Teaching; Psychometry; Quantitative research; Theory of Meaningful Learning.*

RESUMEN

La investigación en educación y enseñanza de las ciencias desempeña un papel clave en la mejora de la calidad de la educación y la comprensión de los retos a los que se enfrentan profesores y estudiantes. Para recopilar datos fiables y relevantes, es esencial que los instrumentos de investigación, como los cuestionarios, estén debidamente validados. El Coeficiente de Validez de Contenido (CVC) es una herramienta importante para evaluar la validez de contenido de un instrumento de recolección de datos, verificando si los ítems que lo componen son claros, pertinentes y relevantes. En el presente trabajo se discute la importancia de la CVC, destacando sus fundamentos y aspectos metodológicos. Por último, se presenta un ejemplo de aplicación

¹ Licenciado em Física, Mestre e Doutor em Educação em Ciências e Professor do curso de Licenciatura em Física da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina/MG – Brasil. E-mail: patrick.fisica@hotmail.com

práctica en la validación de un instrumento que pretende medir el potencial de las clases de Física en la promoción del aprendizaje significativo.

Palabras clave: Enseñanza de Física; Psicometría; Investigación cuantitativa; Teoría del Aprendizaje Significativo.

1. INTRODUÇÃO

Pesquisas em Educação e Ensino de Ciências visam compreender fenômenos relacionados ao ensino, aprendizagem, currículo, formação de professores, entre outros. Busca contribuir para a melhoria do sistema educacional e a promoção de uma alfabetização científica dos estudantes. A coleta de dados é uma parte elementar desse processo, sendo os questionários, frequentemente utilizados como instrumentos de pesquisa. No entanto, a qualidade dos resultados de uma pesquisa depende da validade e confiabilidade dos instrumentos utilizados (Hutz et al., 2015). A etapa de validação de um instrumento é um procedimento essencial para se ter indícios de que ele meça de forma precisa e adequada os construtos de interesse.

Nesse contexto, a validação de questionários emerge como um processo fundamental para assegurar que as questões dos instrumentos de pesquisa sejam relevantes, apropriadas e representativas dos construtos que se pretende medir. A validade de conteúdo, uma das dimensões da validade, é fundamental, uma vez que se relaciona diretamente à representatividade do conteúdo dos itens em relação aos conceitos e temas em estudo.

A psicometria é uma área da psicologia que, entre outras coisas, aborda os procedimentos de elaboração e validação de instrumentos de medida de construtos psicológicos (Hutz et al., 2015). Os procedimentos psicométricos que visam encontrar indícios de validade nos questionários são multifacetados. Segundo os autores, uma das primeiras e mais importantes é a validade de conteúdo. Constata-se essa validade, normalmente, por meio da análise de cada item, mediante um time de juízes, especialistas da área em questão, bem como, por uma amostra do público-alvo a qual o instrumento se destina.

Analisar a opinião dos avaliadores sobre cada um dos itens de um questionário pode se tornar um processo custoso em termos de tempo e recursos humanos. Isso incentivou que, no decorrer das décadas, pesquisadores elaborassem metodologias para que tais procedimentos pudessem ser otimizados e tivessem sua acurácia ampliada (Alexandre; Coluci, 2011). Assim, surgiram diferentes metodologias para aferir concordância entre avaliadores, partindo sempre de uma avaliação dos itens, feita em uma escala de pontos. Com isso, elaboraram-se índices, razões e coeficientes de validade de conteúdo, sempre com vistas a sistematizar e a sintetizar a avaliação de um grupo de pareceristas.

Um deles, o índice de validade de conteúdo (IVC), analisa a proporção ou porcentagem de concordância de avaliadores sobre se o item é adequado ou não ao questionário (Yusoff, 2019). Já a Razão de Validade de Conteúdo (RVC), tem como base a concordância de juízes nos critérios de quanto cada questão é "essencial", "útil, mas não essencial" ou "não necessária" ao instrumento (Tamada; Cunha, 2023). Um terceiro indicador é o Coeficiente de Validade de Conteúdo (CVC), que avalia a concordância dos juízes a partir da análise tridimensional de cada item, mensurando três aspectos: clareza, pertinência e relevância (Reis; Morales, 2020).

Tendo em vista a importância desses indicadores para a sistematização do processo de análise de concordância dos avaliadores, nota-se que as referidas análises são empregadas de maneira substancial nas áreas como administração, psicologia e na saúde, em geral.

Ao mesmo tempo, percebe-se que, na área da educação e ensino de ciências, o uso de tais ferramentas ainda é pouco difundido. Isso motiva a escrita deste artigo a fim de apresentar os fundamentos de uma dessas ferramentas, ao mesmo tempo que se busca demonstrar como implementá-la em um contexto de pesquisa. Para o presente estudo, optou-se pela escolha do CVC, tendo em vista a sua característica de avaliar cada item segundo três parâmetros, sendo então, considerado um indicador de maior amplitude, quando comparado com o IVC e o RVC (Alexandre; Coluci, 2011).

O CVC é uma avaliação quantitativa da validade de conteúdo de itens de um questionário e também do instrumento de pesquisa integral. Ele avalia se as questões elaboradas são relevantes, apropriadas e representativas dos construtos que se pretende medir, através da análise de concordância dos pareceristas. Em outras palavras, o CVC ajuda na análise que visa constatar se o instrumento elaborado, traduzido ou adaptado, mede, de forma adequada, aquilo que se propõe a medir.

O cálculo de concordância do CVC envolve a participação de especialistas no campo de estudo. Essa equipe avalia cada questão do instrumento em relação a diferentes aspectos do seu conteúdo, como clareza, pertinência e relevância. O CVC é então calculado a partir da média aritmética das classificações dadas pelos especialistas para cada item.

Embora o IVC com coeficiente Kappa modificado seja uma técnica mais utilizada em pesquisas recentes, opta-se pelo CVC neste estudo por considerar sua capacidade de analisar cada item do questionário sob três dimensões: clareza, pertinência e relevância. Essa análise tridimensional favorece identificar com maior precisão as fragilidades de cada item e realizar as adaptações necessárias. Acredita-se que o uso do CVC, ainda pouco explorado na área de educação em ciências, pode contribuir significativamente para a qualidade das pesquisas, especialmente no que diz respeito à construção de instrumentos de coleta de dados mais robustos e confiáveis.

A fim de abordar as perspectivas teóricas e práticas do CVC este artigo tem como objetivo geral demonstrar a aplicação do coeficiente na validação de um questionário de pesquisa em educação e ensino de ciências.

De maneira específica o trabalho se propõe a explorar em detalhes o CVC como uma métrica quantitativa utilizada na validação de conteúdo de questionários; descrever um exemplo de validação de conteúdo de um questionário a partir do uso do CVC, apresentando o contexto, os critérios de avaliação estabelecidos e os resultados obtidos; analisar e discutir os resultados obtidos, destacando as implicações do uso do CVC, a identificação de questões problemáticas e o processo de revisão e aprimoramento do questionário; e fornecer diretrizes práticas para pesquisadores e professores interessados em usar o CVC na validação de questionários da área de educação em ciências.

Nota-se, na literatura da área, um número pequeno de estudos que fizeram uso do CVC. Destacam-se as produções de autores como: Silveira e Rocha (2017), que usaram o CVC para validar uma tradução de um instrumento de avaliação do ensino pelos alunos; Araújo (2020), que empregou o CVC na construção e validação de um questionário sobre concepções equivocadas de evolução; Ribeiro e colaboradores (2020), que elaboraram e validaram o conteúdo de uma cartilha sobre diabetes para agentes comunitários de saúde; Ferreira e colaboradores (2022), que usaram o CVC

para elaborar e validar uma avaliação digital de ciências para o Ensino Médio; e Gonçalves e Pedro (2023), que empregaram o CVC para a tradução e validação de uma escala de utilização das tecnologias digitais na gestão escolar.

Tendo em vista a pequena quantidade de trabalhos da área que usam o CVC, ao abordar o seu conceito, seu cálculo, sua aplicação e sua importância, este trabalho visa proporcionar à literatura uma base para pesquisadores que queiram aprimorar a qualidade de seus questionários.

Em um contexto de pesquisa em constante evolução metodológica, em que a educação e ensino de ciências exerce um papel basilar na melhoria do ensino e da aprendizagem de ciências, o uso do CVC como ferramenta de validação de conteúdo mostra-se potente para auxiliar na determinação de indícios de validade de questionários.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. A importância da validade de conteúdo enquanto índice de validação de questionários

A pesquisa em educação e ensino de ciências é um campo multifacetado e dinâmico que busca investigar diferentes aspectos, além de identificar lacunas, desafios e desenvolver estratégias para melhorar a qualidade da educação científica nos diferentes espaços de ensino e aprendizagem (Moreira; Rizzatti, 2020). Para atingir os objetivos de pesquisa, a coleta de dados é uma etapa fundamental. É comum questionários serem escolhidos como instrumentos para esta etapa, devido à sua praticidade e eficiência na coleta de informações de inúmeros participantes. Entretanto, para que os resultados obtidos sejam confiáveis, é fundamental assegurar, entre outras coisas, a qualidade psicométrica dos instrumentos elaborados.

A psicometria estuda a medição e avaliação de aspectos psicológicos, como habilidades, inteligência, personalidade, aprendizagem, ansiedade, entre outros. Nessa área, está incluída as diretrizes sobre a elaboração, tradução e adaptação de testes, análises estatísticas dos dados obtidos e a interpretação dos resultados, a fim de compreender os traços mentais e emocionais dos indivíduos (AERA et al., 2014). Esta área mostra-se imprescindível para a pesquisa em psicologia, avaliação educacional, seleção e gestão de pessoas e em diferentes outros campos os quais a avaliação de atributos psicológicos seja necessário (Hutz et al., 2015).

Para a pesquisa educacional a psicometria mostra-se importante pois contribui com bases teóricas e metodológicas para a construção e validação de instrumentos de medida. A validade de um instrumento refere-se ao grau em que ele mede o que se propõe a medir. A validade de conteúdo, um dos tipos de validade, é fundamental para garantir que os itens do instrumento sejam relevantes, apropriados e representativos do construto em questão.

A validação de um questionário é um processo essencial para determinar se ele realmente avalia os construtos ou conceitos que se propõe a medir no ato de sua elaboração. Uma falha neste processo pode levar a conclusões de pesquisa incorretas e à tomada de decisões inadequadas por quem os utiliza. Dessa maneira, a validação do instrumento é uma etapa crucial que deve ser cuidadosamente conduzida em qualquer pesquisa.

A validade de conteúdo é um tipo de validade analisada em instrumentos de pesquisa. Geralmente, é o primeiro indício de validação buscado. Outros indícios costumam ser analisados de maneira subsequente para consolidação de um instrumento de pesquisa. Alguns deles são a: validade baseada na estrutura interna (que analisa se a teoria que embasou a criação do instrumento converge com o comportamento empírico dos itens), validade convergente (quando o instrumento se correlaciona com outros instrumentos que medem o mesmo construto), validade discriminante (quando o instrumento não se correlaciona com outros instrumentos que medem construtos diferentes), entre outras (Hutz et al., 2015).

Assegurar a existência de indícios de validade de conteúdo de questionários é essencial, uma vez que instrumentos com questões inadequadas ou irrelevantes podem levar a conclusões equivocadas sobre o que se está querendo medir. Isso pode, conseqüentemente, influenciar de forma negativa os resultados de pesquisa. Dessa maneira, defende-se que a avaliação rigorosa das etapas de construção, tradução ou adaptação de um instrumento de coleta de dados é fundamental para assegurar que os resultados da pesquisa sejam confiáveis, válidos e representativos dos construtos em estudo.

2.2. O Coeficiente de Validade de Conteúdo

O CVC é uma métrica desenvolvida por Hernández-Nieto (2002) que visa aferir a validade de conteúdo de um instrumento de pesquisa, como um questionário. Quem avalia cada item é uma equipe de especialistas na área de estudo (o público-alvo também pode usar o CVC para avaliar o instrumento). Julgam-se aspectos sobre o questionário, observando se os itens são relevantes, apropriados e representativos dos construtos que se pretende medir.

Para tanto, os avaliadores especialistas classificam cada item do instrumento com base em critérios predefinidos, que são: clareza da linguagem (o quão compreensível é o item), pertinência prática (se o item é relevante para o instrumento) e relevância teórica (se o item representa o construto que se deseja medir).

Essas classificações são então submetidas a cálculos que resultam no valor médio desses coeficientes para cada item, e também, um valor médio global para o instrumento. De modo a determinar se um item é adequado ou não ao instrumento, a literatura aponta valores mínimos aceitáveis de CVC como critério para a inclusão, revisão ou exclusão de itens do questionário.

O CVC, enquanto ferramenta de validação de conteúdo, pode ser empregado para os mais diferentes fins, como, por exemplo:

a) Desenvolvimento de novos instrumentos de pesquisa – a fim de assegurar que cada item do questionário seja relevante e adequado ao contexto do público-alvo. Pode ser útil ao impedir que questões que não contribuam com a inferência do construto sejam incluídas no questionário.

b) Validade de instrumentos de pesquisa existentes – considerando que é comum fazer uso de instrumentos já presentes na literatura, deve-se atentar que, para um novo contexto específico, é importante verificar se o instrumento se adequa para medir o que se propõe nesse novo contexto ou objetivo de investigação. Assim, o CVC pode servir para conferir validade a tais adaptações.

c) Construção de provas e testes de sala de aula – a fim de garantir que questões de provas e testes estejam mensurando de maneira efetiva os saberes dos alunos para os quais o teste ou prova se destina.

d) Adaptação transcultural de instrumentos – o CVC pode ser usado em pesquisas em que se tem um instrumento em uma língua que não seja a língua materna dos participantes, auxiliando na coerência e validade da tradução, a fim de assegurar que cada item tenha equivalência cultural e linguística.

O procedimento quantitativo empregado no CVC oferece uma abordagem sistemática para a validação de conteúdo. Basicamente, para cada item do instrumento, os avaliadores dão uma nota de 1 a 5 para cada um dos três critérios anteriormente mencionados. Após isso, faz-se uma média da nota de cada critério (para cada item) e subtrai-se, de cada média, um valor referente ao viés de resposta dos avaliadores (conforme poderá ser observado com maiores detalhes na seção seguinte).

O cálculo do CVC, conforme proposto por Hernández-Nieto (2002), utiliza a média simples das notas atribuídas pelos juízes para cada item e cada critério. A média aritmética, como medida de tendência central, oferece uma representação concisa e intuitiva da concordância entre os especialistas, sintetizando as diferentes percepções em um único valor. A escolha pela média simples se justifica por sua facilidade de cálculo e interpretação, além de sua sensibilidade às variações nas notas dos juízes, permitindo captar as nuances na avaliação de cada item (Bussab; Morettin, 2010). Embora a média simples seja uma medida robusta em relação a valores extremos, é importante reconhecer sua sensibilidade a distribuições assimétricas (Barbetta et al., 2024).

Em posse do CVC referente à clareza, à pertinência e à relevância (de cada item), observam-se os valores dos coeficientes avaliados pelos especialistas. Se alguma característica de determinado item for menor que 0,8 a recomendação de Hernández-Nieto (2002) é de adequar sua redação ou removê-lo do instrumento. Por exemplo, se um determinado item for avaliado com coeficiente de 0,85 para Clareza, 0,9 para Pertinência, mas obtiver 0,5 para Relevância, é indicado a sua revisão ou remoção do instrumento.

É possível obter também um CVC para clareza, para pertinência e para relevância considerando o instrumento integral. Para isso, por exemplo, basta calcular a média dos CVC do critério “clareza” de todos os itens, fazendo isso também para pertinência e para relevância. Essa visão geral contribui para verificar em quais dos aspectos o instrumento elaborado (ou traduzido, ou adaptado) necessita de melhorias.

O emprego do CVC não apenas auxilia na confiabilidade e validade dos resultados que serão obtidos pelos questionários, mas também, contribui com os pesquisadores no que tange a economia de tempo e de recursos, possibilitando a otimização da condução de suas investigações.

A fim de apresentar um exemplo de uso do CVC, este artigo exhibirá, nas próximas seções, o processo de validação de conteúdo de um questionário elaborado por Vizzotto (2019) que visa verificar se as aulas de Física do público-alvo apresentam características de aulas que poderiam ou não favorecer o desenvolvimento de uma aprendizagem significativa (Ausubel, 2003). Para isso, acredita-se ser importante uma breve descrição desta teoria, a fim de propiciar uma base teórica para a compreensão dos itens e da estrutura lógica do instrumento validado neste trabalho.

2.3. A Teoria da Aprendizagem Significativa (TAS)

Desenvolvida por David Ausubel (Ausubel, 2003) e muito utilizada em investigações da área de educação e ensino de ciências (Lemos, 2005; Santos, 2007; Valadares, 2011; Moreira, 2012; Darroz, 2018), a TAS é um modelo educacional que se concentra na maneira como os indivíduos adquirem conhecimento e significado a partir da interação com novas informações. A teoria enfatiza a importância de conectar um novo saber com o conhecimento prévio do aprendiz, criando condições favoráveis para tornar a aprendizagem significativa, com maior assimilação e retenção.

Conforme explica o autor, a aprendizagem significativa ocorre quando: o estudante possui um conhecimento prévio relevante (existência de subsunçores); o novo material de aprendizagem é claro e organizado (potencialmente significativo); e quando o aprendiz está ativamente engajado no processo, mantém uma atitude receptiva e estabelece conexões claras entre o novo e o conhecimento prévio (predisposição para aprender) (Ausubel, 2003). Por ser entendida como uma teoria de sala de aula (Moreira, 2012), acredita-se, neste trabalho, ser relevante usar de suas características para conjecturar que tipo de sala de aula teria maior chance de promover uma aprendizagem significativa.

A seção seguinte abordará como o CVC foi empregado na etapa de validação de conteúdo do instrumento que visou mensurar esse fenômeno educacional.

3. METODOLOGIA

3.1. Caracterização

Este trabalho, de característica quantitativa e exploratória (Gil, 2008), pretende mostrar como o CVC pode ser empregado para avaliar a validade de conteúdo de um questionário. Os itens elaborados visam avaliar se as aulas de Física vivenciadas por egressos do Ensino Médio tinham ou não o potencial de promover uma aprendizagem significativa.

O uso do CVC busca assegurar, de maneira sistemática, por meio de uma análise de concordância de especialistas da área, que cada item do questionário tenha clareza, pertinência, relevância para a pesquisa.

3.2. Elaboração de itens

As questões foram elaboradas tendo como base a TAS e obedecendo aos critérios psicométricos de elaboração de itens recomendados por Pasquali (1998). Situações cotidianas de aulas de Física foram representadas por meio de afirmativas. Um primeiro conjunto de itens continham exemplos de aulas que, consoante a TAS, poderia não favorecer uma aprendizagem significativa. O segundo conjunto de itens apresenta situações opostas.

Além da TAS, os 40 itens iniciais do PEF foram elaborados com base em uma análise das dificuldades e desafios enfrentados por alunos no aprendizado de Física (Abud; Maintinguer, 2024). Para a construção dos itens, utilizou-se os seguintes critérios:

- **Situações de Sala de Aula:** Os itens descrevem situações concretas que ocorrem em salas de aula de Física, como a forma de apresentação do conteúdo, as interações entre professor e alunos e o tipo de avaliação utilizada.

- **Características favoráveis e não favoráveis à aprendizagem significativa:** Elaborou-se itens que representam tanto características favoráveis quanto não favoráveis à aprendizagem significativa, de acordo com os princípios da TAS. Por exemplo, o item 7 ("Nas aulas de Física existiam momentos de debate entre professor e alunos, e entre os próprios alunos.") representa uma característica favorável, enquanto o item 21 ("O professor abordava os conteúdos, geralmente fazendo uma exposição teórica para a turma.") representa uma característica não favorável.

O instrumento denominado de "Perfil do Ensino da Física" (PEF) visou refletir os elementos chave da TAS, como:

- **Existência de Subsúncos:** O item 1 ("Na maioria das vezes, o professor iniciava um novo assunto questionando o que os alunos já sabiam sobre o tema.") busca identificar se os professores consideravam o conhecimento prévio dos alunos. A ativação de subsúncos, ou seja, a relação entre o novo conhecimento e o que o aluno já sabe, é essencial para a aprendizagem significativa.
- **Organização do Conteúdo:** Os itens 3 ("Ao ensinar um assunto novo, geralmente o professor abordava aspectos mais gerais do tema para posteriormente detalhar o conceito estudado.") e 4 ("Após as avaliações existiam momentos para refletir sobre os erros e compreender quais eram as suas dificuldades.") buscam verificar se o conteúdo era apresentado de forma organizada e se havia espaço para revisão e reflexão, aspectos que facilitam a compreensão e a assimilação do conhecimento.
- **Predisposição para Aprender:** Os itens 15 ("Sentia-se motivado a aprender física.") e 39 ("A forma com que os assuntos de física eram abordados tornava a aula desmotivadora.") procuram avaliar a motivação dos alunos para aprender física, um fator importante para o engajamento no processo de aprendizagem.

A partir da análise das respostas aos itens do PEF, espera-se obter um "perfil" das aulas de Física, indicando em que medida as práticas pedagógicas relatadas pelos respondentes se aproximam ou se distanciam dos princípios da TAS.

O construto "perfil das aulas de Física", mensurado pelo PEF, refere-se ao conjunto de características e práticas pedagógicas observadas nas aulas de Física, analisadas a partir da ótica da Teoria da Aprendizagem Significativa (TAS). Esse construto busca capturar a dinâmica da sala de aula, considerando a forma como o conteúdo é apresentado, as interações entre professor e alunos, o uso de recursos didáticos e a avaliação da aprendizagem.

As dimensões que compõem o construto "perfil das aulas de Física" são:

- **Organização do Conteúdo:** Clareza na apresentação dos conceitos, sequência lógica na abordagem dos temas, uso de recursos didáticos que facilitem a compreensão.
- **Interação Professor-Aluno:** Diálogo aberto e respeitoso, incentivo à participação dos alunos, feedback construtivo.
- **Estratégias de Ensino:** Utilização de métodos e técnicas de ensino que promovam a aprendizagem significativa, como a resolução de problemas, o trabalho em grupo e a experimentação.
- **Avaliação da Aprendizagem:** Instrumentos de avaliação que verifiquem a compreensão

conceitual e a aplicação dos conhecimentos em diferentes contextos.

Ao mensurar essas dimensões, o PEF busca identificar em que medida as aulas de Física analisadas apresentam características que favorecem ou dificultam a aprendizagem significativa.

Assim, por meio de uma escala Likert, espera-se que o público-alvo (egressos do Ensino Médio) possa responder quais desses itens mais possuem semelhança com a realidade vivenciada por eles durante esta etapa de sua educação.

Itens com características de aulas que poderiam favorecer uma aprendizagem significativa recebem uma pontuação crescente (de 1 a 5), conforme uma escala Likert de 5 pontos (com as alternativas: discordo totalmente, discordo, não concordo nem discordo, concordo e concordo totalmente). Já os itens com as características opostas, possuem também as suas pontuações em numeração contrária (de 5 a 1).

A escala Likert de 5 pontos foi escolhida para o instrumento PEF por ser amplamente utilizada em pesquisas na área de educação (Trojan; Sipraki, 2015) e por permitir a mensuração de atitudes e opiniões em um continuum, capturando diferentes níveis de concordância ou discordância (Monte, 2020).

O somatório de pontos da escala procura representar o construto denominado de "perfil das aulas de Física" desses participantes.

3.3. Envio para especialistas da área

Os itens, após o processo de elaboração, foram enviados por e-mail para 3 especialistas da área, professores de Física, que voluntariamente aceitaram atuar como avaliadores das questões, contribuindo com suas opiniões e pareceres, pontuando cada item segundo os três critérios mencionados e sugerindo melhorias na sua escrita, quando necessário. Os critérios de inclusão dos avaliadores foram:

- **Formação Acadêmica:** Possuir doutorado em Ensino de Física ou área afim.
- **Experiência Docente:** Ter experiência mínima de 5 anos no Ensino Médio.
- **Produção Científica:** Possuir publicações na área de Ensino de Física, com ênfase em temas relacionados à TAS.

A expertise dos avaliadores foi verificada por meio da análise de seus currículos Lattes e de suas publicações.

O tempo de espera até o retorno de todas as avaliações foi de 2 meses.

3.4. Critérios de avaliação e cálculo do CVC

Para chegar ao coeficiente, em um primeiro momento, deve-se calcular a média de cada critério, para cada item, onde a "Média = Soma das pontuações dos juízes / número de juízes":

$$\mu = \frac{\Sigma \text{pontuação}}{n \text{ juízes}}$$

Para o cálculo do CVC, divide-se a média pelo maior valor de escore possível na avaliação (ou seja, 5 pontos), e subtrai-se desse resultado um valor referente a uma correção de viés dos juízes, conforme pode ser visualizado:

$$CVC = \left(\left[\frac{\mu}{5} \right] - \text{fator de correção de viés} \right)$$

O viés de juízes consiste em uma influência subjetiva que os avaliadores podem exercer no momento da análise dos itens. Isso pode acontecer quando os especialistas interpretam as questões do instrumento de forma distinta, em virtude de suas próprias concepções, vivências, preconceitos, ou perspectivas individuais. Esse viés é capaz de influenciar a coerência da validade e confiabilidade dos resultados de um instrumento. Para mitigá-lo, recomenda-se selecionar os juízes de maneira cuidadosa e apresentar instruções claras sobre como a avaliação deve ser feita. Além disso, é possível realizar correções estatísticas nos dados já prevendo a existência desse fenômeno, motivo pelo qual Hernández-Nieto (2002) propôs a correção usada no CVC.

Este fator é calculado da seguinte maneira:

$$\text{Fator de correção de viés} = \left(\frac{1}{\text{número de juízes}} \right)^{\text{número de juízes}}$$

Por exemplo, caso 2 juízes façam a avaliação dos itens, o viés será:

$$\text{Viés} = \left(\frac{1}{2} \right)^2 = 0,5^2 = 0,25$$

No entanto, se o número de juízes for de 3 pessoas, o viés será de:

$$\text{Viés} = \left(\frac{1}{3} \right)^3 = 0,333^3 = 0,037$$

É possível notar que o viés é considerado grande ao ter apenas dois juízes, mas, com o acréscimo de somente mais um juiz, o valor de viés diminui significativamente. Isso justifica a decisão de sempre ter a contribuição de pelo menos 3 especialistas da área para a avaliação dos itens.

É importante destacar que a estimação de viés de juízes é um processo complexo, que vai além da simples consideração de vieses individuais. Diversas fontes de viés podem estar presentes na avaliação, incluindo vieses sistemáticos, como o viés de leniência e o viés de severidade, além da influência de fatores como a experiência dos juízes, a familiaridade com o tema e a compreensão das instruções (Leonardo, 2013).

A fórmula de correção de viés utilizada no CVC leva em consideração o número de juízes, o que a torna mais precisa em estudos com diferentes tamanhos de equipe de avaliação. No entanto, é fundamental reconhecer as limitações do termo de correção de viés. O método se baseia em pressupostos que podem não eliminar completamente a influência da subjetividade dos juízes. Mesmo após a correção, ainda há espaço para variações individuais na avaliação. A escolha do método de correção de viés proposto por Hernández-Nieto (2002) para este estudo se justifica por sua adequação ao tamanho da equipe de avaliação e por sua capacidade de minimizar a influência de diferentes fontes de viés na análise dos resultados.

É fundamental reconhecer, no entanto, que o fator de correção não elimina os vieses inerentes a cada avaliador individualmente. Ele atua como um ajuste estatístico que incentiva a participação de pelo menos três especialistas, mas não garante a isenção de vieses.

Conforme referido, para minimizar a influência dos vieses, além da inclusão de pelo menos três avaliadores, recomenda-se adotar estratégias como seleção criteriosa dos especialistas, instruções claras e detalhadas, treinamento dos avaliadores, cegamento dos avaliadores e análise da consistência das avaliações. Acredita-se que a combinação dessas medidas contribui para a obtenção de uma avaliação mais objetiva e representativa do construto em questão.

3.5. Interpretação dos resultados e revisão dos itens

Essa etapa se centrou em identificar itens que estivessem abaixo do ponto de corte em alguma das características avaliadas. O coeficiente mínimo, definido por Hernández-Nieto (2002), é de CVC = 0,8. Itens com pontuação abaixo de 0,8 em qualquer um dos critérios foram cuidadosamente revisados, com base nas sugestões dos avaliadores. As modificações realizadas visaram aprimorar a clareza da linguagem, a pertinência prática e a relevância teórica dos itens, buscando atender aos critérios de qualidade estabelecidos.

Após identificar as questões consideradas inadequadas ou sujeitas a exclusão, estes itens foram revisados qualitativamente a partir da teoria a qual foram elaborados, de modo a adequar a sua construção. Quando observado que, mesmo com as adequações, os itens não estariam aptos a compor o instrumento, atendeu-se a recomendação de não os manter na versão final do questionário.

Em seguida a este procedimento de revisão e/ou reformulação, os itens alterados foram enviados novamente para os especialistas, tendo em vista uma segunda rodada de avaliação. Assim, com as novas notas para esses itens, foi possível calcular um novo CVC para eles e para o instrumento integral.

Na sequência, apresentam-se os resultados obtidos a partir do retorno das avaliações dos especialistas.

4. OS RESULTADOS E DISCUSSÕES

4.1. Etapa 1 – Os itens elaborados

Após o retorno da primeira rodada de avaliação de itens, as notas foram computadas em uma planilha eletrônica e suas médias calculadas. Em posse das mesmas, chegou-se ao CVC de cada item, para cada característica analisada. Apresenta-se os resultados das análises realizadas, iniciando pela descrição de cada item elaborado, conforme os quadros 1 e 2.

Quadro 1 - Itens com características de aulas que poderiam favorecer uma aprendizagem significativa

- 1- O professor iniciava um novo assunto sempre questionando o que os alunos já sabiam sobre o tema.
- 2- A explicação dos conceitos de física geralmente relacionava-se com situações observadas no cotidiano dos alunos.
- 3- O professor organizava as ideias, conceitos e informações de forma esquematizada através de mapas conceituais.
- 4- Ao ensinar um assunto novo, geralmente o professor abordava aspectos mais gerais do tema para

posteriormente detalhar o conceito estudado.

5- Era comum a realização de estudos e exercícios em grupo.

6- O professor estimulava os alunos a buscarem ampliar seus conhecimentos através de pesquisas, vídeos, reportagens e simulações.

7- Nas aulas de Física existiam momentos de debate entre professor e alunos, e entre os próprios alunos.

8- Em alguns assuntos trabalhados, havia a realização de atividades experimentais, nas quais os alunos manuseavam os experimentos.

9- Durante as avaliações, os conteúdos eram abordados em situações diferentes daquelas usadas como exemplo no momento em que foram ensinadas.

10- Durante as aulas havia oportunidades para o professor ouvir as compreensões dos alunos sobre o assunto que estava sendo abordado.

11- Uma parte da avaliação era composta pela Autoavaliação.

12- Após as avaliações existiam momentos para refletir sobre os erros e compreender quais eram as minhas dificuldades.

13- Após ter aula sobre um novo conceito físico, geralmente eu conseguia observar tais fenômenos em situações do meu cotidiano.

14- Após a conclusão do Ensino Médio eu conseguia compreender os fenômenos físicos presentes no meu cotidiano de forma mais completa.

15- Me sentia motivado a assistir às aulas de Física.

16- Me sentia motivado a aprender física.

17- Ao ensinar um novo conteúdo, o professor apresentava o contexto histórico de onde derivou aquele conhecimento.

18- Geralmente o professor apresentava algum material introdutório (textos científicos, reportagens, vídeos, filmes, simulações) antes de iniciar o estudo de um novo conceito, a fim de familiarizar os alunos o tema.

19- Acontecimentos do cotidiano eram discutidos nas aulas, quando estes tinham relação com os conteúdos estudados.

20- Era comum o professor relacionar os assuntos de Física com outras disciplinas escolares.

Fonte: autor.

Na sequência, apresenta-se o quadro 2 com itens de características de aulas que poderiam não favorecer uma aprendizagem significativa.

Quadro 2 - Itens com características de aulas que poderiam não favorecer uma aprendizagem significativa

21- O professor abordava os conteúdos, geralmente fazendo uma exposição teórica para a turma.

22- A avaliação acontecia por meio de provas escritas.

23- Em alguns assuntos trabalhados, havia a realização de atividades experimentais nas quais, o professor apenas, manuseava os experimentos.

24- As questões das provas eram semelhantes às dos exercícios e exemplos trabalhados nas aulas.

25- Durante as aulas, o professor prezava pelo silêncio da turma a fim de prestarem atenção na explicação transmitida.

26- As aulas de física tinham como roteiro um livro didático ou uma apostila.

27- Após o estudo de um novo assunto, geralmente o professor trabalhava uma lista de exercícios para fixação do conteúdo.

28- Os exercícios realizados eram parecidos com os exemplos abordados pelo professor durante a explicação do conteúdo.

29- A relação entre professor e alunos era distante e formal.

30- Os estudos e exercícios eram realizados normalmente de forma individual.

- 31- Geralmente as questões de provas e exercícios eram consideradas apenas como certas ou erradas, ignorando o raciocínio dos alunos.
- 32- O desenvolvimento de um exercício deveria seguir o modelo ensinado pelo professor.
- 33- O professor detinha o conhecimento da disciplina e o transmitia aos alunos.
- 34- A sala de aula era disposta em fileiras de classes.
- 35- Não existiam momentos para debates entre professor e alunos e entre os próprios alunos.
- 36- A avaliação buscava observar se o aluno memorizou a matéria ensinada pelo professor.
- 37- Era grande o número de fórmulas que se necessitava memorizar.
- 38- Normalmente os conceitos físicos estudados eram esquecidos logo depois da prova.
- 39- A forma com que os assuntos de física eram abordados tornava a aula desmotivadora.
- 40- Os conteúdos de física tinham pouca relação com situações do cotidiano.

Fonte: autor.

Destaca-se, no que tange a elaboração dos itens, que foram consideradas situações referentes ao contexto da sala de aula, interação com o docente e colegas e opinião referente à sensação de aprendizagem de Física.

4.2. Etapa 2 – As notas dos juízes

O quadro 3, apresenta todas as notas dadas pelos três juízes para cada item, em cada um dos critérios analisados.

Quadro 3 - Nota dos juízes para itens com características que poderiam favorecer uma aprendizagem significativa

Item	Clareza			Pertinência			Relevância		
	Juiz 1	Juiz 2	Juiz 3	Juiz 1	Juiz 2	Juiz 3	Juiz 1	Juiz 2	Juiz 3
1	3	4	4	4	4	5	5	4	4
2	4	4	5	2	3	1	2	2	2
3	3	3	2	4	5	4	2	2	2
4	4	4	5	5	4	4	5	5	4
5	5	5	4	3	1	2	2	2	2
6	5	5	4	4	4	5	4	5	4
7	4	5	5	5	4	5	4	5	5
8	3	3	2	5	4	4	5	4	4
9	4	4	5	2	1	2	2	2	2
10	3	3	2	4	5	4	4	4	5
11	5	5	5	4	5	4	4	4	5
12	4	4	5	5	4	4	4	5	4
13	3	3	2	5	4	4	4	4	5
14	3	3	3	5	5	5	4	5	4
15	4	4	5	3	2	2	2	2	2
16	3	2	2	5	4	4	5	5	4
17	4	4	5	5	4	5	4	5	4

18	4	5	4	5	5	4	5	5	4
19	4	4	5	3	4	2	2	2	2
20	5	5	4	4	4	5	4	5	5

Fonte: autor.

Observa-se que, para este primeiro conjunto de itens, as notas dos critérios “Clareza” e “Relevância” foram mais altas do que para o critério “Pertinência”. Na sequência, o quadro 4 exibe-se as notas do segundo conjunto de itens.

Quadro 4 - Nota dos juízes para itens com características que poderiam não favorecer uma aprendizagem significativa

Item	Clareza			Pertinência			Relevância		
	Juiz 1	Juiz 2	Juiz 3	Juiz 1	Juiz 2	Juiz 3	Juiz 1	Juiz 2	Juiz 3
21	3	3	2	5	4	4	2	2	2
22	4	4	5	2	3	4	3	4	5
23	3	3	4	2	3	3	5	5	4
24	5	5	5	3	2	3	4	3	4
25	5	3	4	5	3	4	4	3	4
26	3	4	5	2	4	3	4	3	4
27	3	3	5	5	3	2	4	3	4
28	5	4	3	2	1	3	4	3	4
29	4	5	3	1	3	3	4	3	4
30	5	3	3	3	3	5	4	5	5
31	3	2	4	5	4	5	5	4	5
32	3	4	4	3	2	3	5	5	4
33	4	5	4	2	3	4	5	4	3
34	5	5	5	3	4	2	5	3	3
35	5	5	4	4	5	4	5	5	4
36	5	4	4	3	2	2	4	4	4
37	3	2	4	5	4	5	5	4	4
38	5	5	4	5	4	5	5	4	4
39	5	5	4	5	4	5	5	4	4
40	4	4	5	5	5	5	4	4	5

Fonte: autor.

Para este segundo conjunto de itens, em média, as notas de “Clareza” foram mais baixas. Em uma primeira análise, é possível afirmar que há maior quantidade de itens que possuirão CVC abaixo do ponto de corte. Apresenta-se, na seção seguinte, os resultados do cálculo dos coeficientes para cada questão.

4.3. Etapa 3 – Determinação do CVC

4.3.1. Cálculo do viés e do coeficiente

Conforme observado na seção anterior o fator de correção do viés tendo 3 juízes resulta no valor de 0,037 (valor que deve ser subtraído do CVC de cada item, para cada critério analisado).

No quadro 5, exibe-se os CVC dos itens com características de aulas que poderiam favorecer uma aprendizagem significativa.

Quadro 5 - CVC dos itens com características que poderiam favorecer uma aprendizagem significativa

Item	CVC_Clareza	CVC_Pert	CVC_Rel	Item	CVC_Clareza	CVC_Pert	CVC_Rel
1	0,696	0,83	0,83	11	0,963	0,83	0,83
2	0,83	0,363	0,363	12	0,83	0,83	0,83
3	0,496	0,83	0,363	13	0,496	0,83	0,83
4	0,83	0,83	0,896	14	0,563	0,963	0,83
5	0,896	0,363	0,363	15	0,83	0,43	0,363
6	0,896	0,83	0,83	16	0,43	0,83	0,896
7	0,896	0,896	0,896	17	0,83	0,896	0,83
8	0,496	0,83	0,83	18	0,83	0,896	0,896
9	0,83	0,296	0,363	19	0,83	0,563	0,363
10	0,496	0,83	0,83	20	0,896	0,83	0,896

Fonte: autor.

No quadro 5, nota-se que apenas 8 itens foram considerados adequados em todos os três critérios (itens 4, 6, 7, 11, 12, 17, 18 e 20). Ou seja, os demais, necessitam passar por adequações, se sugeridas pelos juízes, ou serem excluídos do instrumento.

Ao analisar os CVC desse primeiro conjunto de itens, observa-se que, é o critério “Clareza” que aponta menores valores de coeficiente. Isso significa que, segundo a concordância dos avaliadores, grande parte dos itens elaborados podem ser considerados importantes para medir o constructo desejado (pertinência), possuem coerência à teoria que os embasa (relevância), mas necessitam de adequações na sua escrita/apresentação (clareza).

O quadro 6 apresenta os CVC para o segundo grupo de itens elaborados.

Quadro 6 - CVC dos itens com características que poderiam não favorecer uma aprendizagem significativa

Item	CVC_Clareza	CVC_Pert	CVC_Rel	Item	CVC_Clareza	CVC_Pert	CVC_Rel
21	0,496	0,83	0,363	31	0,563	0,896	0,896
22	0,83	0,563	0,763	32	0,696	0,496	0,896
23	0,63	0,496	0,896	33	0,83	0,563	0,763
24	0,963	0,496	0,696	34	0,963	0,563	0,696
25	0,763	0,763	0,696	35	0,896	0,83	0,896
26	0,763	0,563	0,696	36	0,83	0,43	0,763
27	0,696	0,63	0,696	37	0,563	0,896	0,83

28	0,763	0,363	0,696	38	0,896	0,896	0,83
29	0,763	0,43	0,696	39	0,896	0,896	0,83
30	0,696	0,696	0,896	40	0,83	0,963	0,83

Fonte: autor.

Ao analisar o Quadro 6, nota-se que, para este conjunto de itens, apenas 4 deles foram considerados adequados para compor o instrumento final em todos os três critérios analisados (itens 35, 38, 39 e 40).

Por fim, foi calculado o CVC de cada critério avaliado para o instrumento integral, desta primeira rodada avaliativa, ou seja, considerando todas as 40 questões elaboradas. Os valores foram os seguintes: Clareza: 0,756; Pertinência: 0,701; Relevância: 0,736. Isso significa que não há possibilidade de manter todos os 40 itens no instrumento final, pois, em conjunto, suas notas não contribuem para o CVC geral passar do ponto de corte recomendado pela literatura (CVC \geq 0,8). Ou seja, deve ser feito um refinamento dos itens. Revisar aqueles que tiveram sugestões de melhoria, reavaliá-los e calcular um novo CVC individual (para os itens reformulados) e integral (para o instrumento).

4.3.2. Interpretação do CVC

Para facilitar a visualização de quais itens foram ou não considerados aceitáveis, os quadros 7 e 8 apresentam a interpretação do CVC, com base no critério do ponto de corte. Ou seja, conforme referido, acima de 0,8 o item é considerado "aceitável" e abaixo de 0,8, inaceitável.

Para a questão ser selecionada para o instrumento integral, necessita obter "aceitável" em cada um dos três critérios. O quadro 7 apresenta a interpretação dos CVC para os itens com características de aulas que poderiam favorecer uma aprendizagem significativa.

Quadro 7 - Interpretação do CVC para os itens com características que poderiam favorecer uma aprendizagem significativa

Item	Interpretação			Item	Interpretação		
	Clareza	Pertinência	Relevância		Clareza	Pertinência	Relevância
1	Inaceitável	Aceitável	Aceitável	11	Aceitável	Aceitável	Aceitável
2	Aceitável	Inaceitável	Inaceitável	12	Aceitável	Aceitável	Aceitável
3	Inaceitável	Aceitável	Inaceitável	13	Inaceitável	Aceitável	Aceitável
4	Aceitável	Aceitável	Aceitável	14	Inaceitável	Aceitável	Aceitável
5	Aceitável	Inaceitável	Inaceitável	15	Aceitável	Inaceitável	Inaceitável
6	Aceitável	Aceitável	Aceitável	16	Inaceitável	Aceitável	Aceitável
7	Aceitável	Aceitável	Aceitável	17	Aceitável	Aceitável	Aceitável
8	Inaceitável	Aceitável	Aceitável	18	Aceitável	Aceitável	Aceitável
9	Aceitável	Inaceitável	Inaceitável	19	Aceitável	Inaceitável	Inaceitável
10	Inaceitável	Aceitável	Aceitável	20	Aceitável	Aceitável	Aceitável

Fonte: autor.

Neste primeiro conjunto de itens, observa-se que as questões 4, 6, 7, 11, 12, 17, 18, e 20 foram consideradas aceitáveis em todos os critérios de avaliação. Se algum item for classificado como

“inaceitável” nos três critérios, recomenda-se sua exclusão. Já os itens com apenas uma classificação “inaceitável”, exceto no critério “relevância”, podem ser reelaborados e reavaliados.

O quadro 8 apresenta a interpretação do CVC para os itens com características de aulas que poderiam não favorecer uma aprendizagem significativa.

Quadro 8 - Interpretação do CVC para os itens com características que poderiam não favorecer uma aprendizagem significativa

Item	Interpretação			Item	Interpretação		
	Clareza	Pertinência	Relevância		Clareza	Pertinência	Relevância
21	Inaceitável	Aceitável	Inaceitável	31	Inaceitável	Aceitável	Aceitável
22	Aceitável	Inaceitável	Inaceitável	32	Inaceitável	Inaceitável	Aceitável
23	Inaceitável	Inaceitável	Aceitável	33	Aceitável	Inaceitável	Inaceitável
24	Aceitável	Inaceitável	Inaceitável	34	Aceitável	Inaceitável	Inaceitável
25	Inaceitável	Inaceitável	Inaceitável	35	Aceitável	Aceitável	Aceitável
26	Inaceitável	Inaceitável	Inaceitável	36	Aceitável	Inaceitável	Inaceitável
27	Inaceitável	Inaceitável	Inaceitável	37	Inaceitável	Aceitável	Aceitável
28	Inaceitável	Inaceitável	Inaceitável	38	Aceitável	Aceitável	Aceitável
29	Inaceitável	Inaceitável	Inaceitável	39	Aceitável	Aceitável	Aceitável
30	Inaceitável	Inaceitável	Aceitável	40	Aceitável	Aceitável	Aceitável

Fonte: autor.

Para este segundo conjunto de questões, pode-se notar que 5 delas foram classificadas como inaceitáveis em todos os critérios (25, 26, 27, 28 e 29), sendo então, eliminadas do instrumento. Do restante, 11 itens foram revisados e/ou eliminados e os outros 4, incluídos no questionário final.

4.4. Revisão dos itens

Essa etapa consistiu na análise de qual critério as questões apresentaram fragilidade. No momento da primeira rodada de avaliação solicitou-se que os juízes apresentassem sugestões de como o item poderia ser melhorado. Assim, 8 deles tiveram sugestões de melhorias. Todas as recomendações dos juízes foram atendidas.

Após a readequação, os itens foram enviados novamente aos avaliadores para uma segunda rodada de avaliação. O quadro 9 apresenta as notas dos itens que não foram eliminados em etapas anteriores, incluindo as novas notas das 8 questões revisadas.

Quadro 9 – Notas após a segunda rodada de avaliação dos itens revisados

Item	Clareza			Pertinência			Relevância		
	<u>Juiz 1</u>	<u>Juiz 2</u>	<u>Juiz 3</u>	<u>Juiz 1</u>	<u>Juiz 2</u>	<u>Juiz 3</u>	<u>Juiz 1</u>	<u>Juiz 2</u>	<u>Juiz 3</u>
1	5	4	5	4	4	5	5	4	4
4	4	4	5	5	4	4	5	5	4
6	5	5	4	4	4	5	4	5	4
7	4	5	5	5	4	5	4	5	5
8	5	4	4	5	4	4	5	4	4
10	4	5	4	4	5	4	4	4	5

11	5	5	5	4	5	4	4	4	5
12	4	4	5	5	4	4	4	5	4
13	5	4	4	5	4	4	4	4	5
14	4	4	5	5	5	5	4	5	4
16	4	5	4	5	4	4	5	5	4
17	4	4	5	5	4	5	4	5	4
18	4	5	4	5	5	4	5	5	4
20	5	5	4	4	4	5	4	5	5
31	5	4	4	5	4	5	5	4	5
35	5	5	4	4	5	4	5	5	4
37	5	4	4	5	4	5	5	4	4
38	5	5	4	5	4	5	5	4	4
39	5	5	4	5	4	5	5	4	4
40	4	4	5	5	5	5	4	4	5

Fonte: autor.

Em posse desse quantitativo, foi possível calcular novamente o CVC dos itens revisados, conforme pode ser observado no quadro 10.

Quadro 10 - CVC após a segunda rodada de avaliação dos itens revisados

Item	CVC_Clareza	CVC_Pertinência	CVC Relevância
1	0,896	0,830	0,830
4	0,830	0,830	0,896
6	0,896	0,830	0,830
7	0,896	0,896	0,896
8	0,830	0,830	0,830
10	0,830	0,830	0,830
11	0,963	0,830	0,830
12	0,830	0,830	0,830
13	0,830	0,830	0,830
14	0,830	0,963	0,830
16	0,830	0,830	0,896
17	0,830	0,896	0,830
18	0,830	0,896	0,896
20	0,896	0,830	0,896
31	0,830	0,896	0,896
35	0,896	0,830	0,896
37	0,830	0,896	0,830
38	0,896	0,896	0,830
39	0,896	0,896	0,830
40	0,830	0,963	0,830

Fonte: autor.

Nota-se que, após o processo de refinamento, revisão e reavaliação, os CVC dos 20 itens restantes foram todos superiores ao ponto de corte, para todos os critérios de análise. Ou seja, os 8 itens revisados foram considerados aptos a compor o questionário. Isso indica que esses 20 itens, a partir dessa ferramenta e desse conjunto de três avaliadores, pode ser considerado com indícios de validade de conteúdo, sendo assim, recomendável que integrem o instrumento final.

Por sua vez, os demais itens foram excluídos (2, 3, 5, 9, 15, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34 e 36).

Finalmente, foi possível calcular novamente o CVC para o instrumento integral, chegando-se a nota 0,86 para "Clareza", 0,87 para "Pertinência", e 0,85 para Relevância, todos os CVC acima do ponto de corte, indicando boa qualidade do instrumento em termos de validade de conteúdo.

Os itens que apresentaram os maiores índices de CVC foram aqueles que descreviam situações de sala de aula claramente alinhadas ou contrárias aos princípios da TAS, como o uso de recursos didáticos variados, a interação entre professor e alunos e a avaliação da compreensão conceitual. Os itens que apresentaram os menores índices de CVC, embora ainda acima do ponto de corte, foram aqueles que descreviam situações mais complexas ou que exigiam maior interpretação por parte dos avaliadores. Esses resultados sugerem que, no que tange o conteúdo do instrumento, o PEF é capaz de diferenciar as características favoráveis e não favoráveis à aprendizagem significativa, de acordo com a TAS.

O quadro 11 apresenta os itens do instrumento final.

Quadro 11 – Itens do instrumento final com indícios de validade de conteúdo

Após as avaliações existiam momentos para refletir sobre os erros e compreender quais eram as suas dificuldades.

Sentia-se motivado a aprender física.

Era grande o número de fórmulas que se necessitava memorizar sem fazer uma análise do seu significado

Após a conclusão do Ensino Médio você conseguia compreender os fenômenos físicos presentes no seu cotidiano de forma mais completa.

Na maioria das vezes, o professor iniciava um novo assunto questionando o que os alunos já sabiam sobre o tema.

Não existiam momentos para debates entre professor e alunos e entre os próprios alunos.

Após ter aula sobre um novo conceito físico, geralmente você conseguia observar tais fenômenos em situações do seu cotidiano.

Ao ensinar um novo conteúdo, o professor apresentava o contexto histórico de onde derivou aquele conhecimento.

Geralmente as questões de provas e exercícios eram consideradas apenas como certas ou erradas, ignorando ou desconsiderando o desenvolvimento da resolução da mesma.

Nas aulas de Física havia a realização de atividades experimentais, nas quais os alunos manuseavam os experimentos.

Era comum o professor relacionar os assuntos de Física com outras disciplinas escolares.

A forma com que os assuntos de física eram abordados tornava a aula desmotivadora.

Nas aulas de Física existiam momentos de debate entre professor e alunos, e entre os próprios alunos.

Uma parte da avaliação era composta pela Autoavaliação.

Normalmente os conceitos físicos estudados eram esquecidos logo depois da prova.

O professor estimulava os alunos a buscarem ampliar seus conhecimentos através de pesquisas, vídeos, reportagens e simulações.

Geralmente o professor apresentava algum material introdutório (textos científicos, reportagens, vídeos, filmes, simulações) antes de iniciar o estudo de um novo conceito, a fim de familiarizar os alunos com o tema.

Os conteúdos de física tinham pouca relação com situações do cotidiano.

Durante a aula o professor permitia que os alunos falassem sobre o assunto que estava sendo abordado.

Ao ensinar um assunto novo, geralmente o professor abordava aspectos mais gerais do tema para posteriormente detalhar o conceito estudado.

Fonte: autor.

É importante destacar que a exclusão de itens inadequados é comum e faz parte do processo de busca por indícios de validade. Considerando isso, faz-se importante que pesquisadores sempre elaborem um número superior de itens, tendo em vista o número de questões que se deseja ter no instrumento final. Neste caso particular, foi elaborado o dobro de itens.

Outro ponto relevante de destacar é a importância de misturar as questões na disposição que serão apresentadas no instrumento. Ou seja, é importante embaralhar itens com características favoráveis e não favoráveis. Isso dificulta com que o respondente crie padrões de resposta baseado no tipo de afirmação que está lendo (favorável ou não favorável).

Por fim, salienta-se que, conforme abordado na fundamentação teórica, a validação de conteúdo é uma das maneiras de se buscar indícios de validade para um instrumento. Geralmente é a primeira. No entanto, a busca por outras validades é fundamental para a consolidação de um questionário. Uma delas é a validade baseada na estrutura interna, que exige, após a validade de conteúdo e de uma primeira versão final do instrumento, um "teste piloto", seguido de uma análise fatorial exploratória e/ou confirmatória (Hutz et al., 2015) a fim de verificar se o comportamento empírico dos itens está em consonância com a estrutura teórica na qual ele foi elaborado.

5. CONSIDERAÇÕES FINAIS

A validação de questionários desempenha um papel essencial na pesquisa em educação e ensino de ciências, certificando que os instrumentos de coleta de dados utilizados tenham indícios de que são confiáveis, válidos e representativos dos construtos que se pretende medir.

Este artigo buscou demonstrar a aplicação do CVC na validação de um questionário de pesquisa. Para esse fim, usou-se do CVC para analisar o conteúdo de itens elaborados para criar um questionário que buscasse investigar se as aulas de Física que o público-alvo havia participado em suas vidas possuía características de aulas que poderiam favorecer ou não uma aprendizagem significativa.

De um número inicial de 40 itens, o processo de avaliação considerou 28 itens inadequados. 8 deles foram revisados, reelaborados, reavaliados e aprovados, e outros 20, excluídos do instrumento final.

Considerou-se atingido objetivo geral deste manuscrito ao apresentar o CVC, seus fundamentos, sua importância e um exemplo de como empregá-lo na avaliação de questões elaboradas para fins de pesquisa.

A partir das métricas do CVC foi possível:

- Notar questões inadequadas que, se fossem mantidas no instrumento final, não atenderiam aos critérios necessários para caracterizar um instrumento válido. Isso destaca a precisão da ferramenta na detecção de questões problemáticas.
- Revisar itens que, em algum dos critérios, não foram considerados adequados, mas que, com melhorias na sua redação, puderam ser considerados aptos a se manter no instrumento final. O CVC auxiliou na finalização de um instrumento com questões de maior relevância e representatividade do constructo que se desejou mensurar.
- Aumentar os indícios de validade de conteúdo do questionário. Ao comparar o CVC geral antes e depois da revisão dos itens, notou-se que, as questões que compuseram o instrumento final, apresentaram um aumento no CVC médio de cada critério analisado. Isso representa um indício de maior intensidade desse tipo de validade psicométrica.

Ao mesmo tempo em que o CVC se mostra uma ferramenta importante para pesquisadores que fazem suas investigações com questionário, é relevante constatar que o seu emprego não está desprovido de desafios. Um deles é que a ferramenta tem como base a análise pessoal de um time de avaliadores. Isso pode acarretar inclusão de variações subjetivas nos resultados, motivo pelo qual se recomenda a escolha cuidadosa do time de especialistas que avaliará o estudo.

Este artigo ressalta a relevância do uso de ferramentas como o CVC para fomentar a construção de instrumentos de coleta de dados válidos, sendo este, o primeiro passo a ser cumprido para que os resultados de pesquisa se aproximem do fenômeno empírico que se deseja investigar.

O uso do CVC não apenas auxilia na observação de indícios de validade do questionário, mas contribui também para que pesquisadores possam economizar tempo, recursos humanos e materiais, o que pode ajudar no êxito de uma pesquisa, na totalidade. Isto posto, o uso do CVC como ferramenta integrante da metodologia de validação de questionários é fortemente aconselhado para pesquisadores de nossa área.

6. REFERÊNCIAS

ABUD, Renata Morri Perroni; MAINTINGUER, Sandra Imaculada. As dificuldades e os desafios do ensino de física (para o ensino médio e novo ensino médio) no contexto pós pandemia. **Infinitum: Revista Multidisciplinar**, v. 7, n. 12, p. 4–20, 2024

ALEXANDRE, Neusa Maria Costa; COLUCI, Marina Zambon Orpinelli. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Ciência & Saúde Coletiva**, v. 16, n. 1, p. 3061-3068, 2011.

American Educational Research Association (AERA); American Psychological Association (APA); National Council On Measurement In Education (NCME). **Standards for educational and psychological testing**. Washington, DC: APA, 2014.

ARAÚJO, Leonardo Augusto Luvison. Concepções equivocadas sobre evolução biológica: um estudo comparativo entre graduandos em ciências biológicas e pós-graduandos. **Investigações em Ensino de Ciências**, v. 25, n. 2, p. 332-346, 2020.

AUSUBEL, David Paul. **Aquisição e retenção de conhecimentos**: uma perspectiva cognitiva. Lisboa: Plátano, 2003. 405 p.

BARBETTA, Pedro Alberto; REIS, Marcelo Menezes; BORNIA, Antonio Cezar. **Estatística:** para cursos de engenharia e informática. 4ª edição. São Paulo: Atlas, 2024.

BUSSAB, Wilton de Oliveira; MORETTIN, Pedro Alberto. **Estatística básica.** 6a edição. São Paulo: Saraiva, 2010, 540 p.

DARROZ, Luiz Marcelo. Aprendizagem significativa: a teoria de David Ausubel. **Revista Espaço Pedagógico**, v. 25, n. 2, p. 576-580, 2018.

FERREIRA, Marcelo; NOGUEIRA, Danielle; SILVA FILHO, Olavo; COSTA, Marcos; SOARES NETO, Joaquim. A WebQuest como proposta de avaliação digital no contexto da aprendizagem significativa crítica em ciências para o ensino médio. **Revista Pesquisa e Debate em Educação**, v. 12, n. 1, p. 1-32, 2022.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social.** 6. ed. Barueri: Atlas, 2008.

GONÇALVES, Antoniel; PEDRO, Neusa. Adaptação e validação do conteúdo da escala de utilização das tecnologias digitais na gestão escolar. **Research, Society and Development**, v. 12, n. 4, p. 1-15, 2023.

HERNÁNDEZ-NIETO, Rafael. **Contributions to statistical analysis:** The Coefficients of Proportional Variance, Content Validity and Kappa. Mérida: Universidad de Los Andes, v. 193, 2002.

HUTZ, Cláudio Simon; BANDEIRA, Denise Rushel; TRENTINI, Clarissa Marcelli. **Psicometria.** Porto Alegre: Artmed, 2015.

LEMOS, Evelyse dos Santos. Situando a teoria de aprendizagem significativa na prática docente, na formação de professores e nas investigações educativas em ciências. **Revista Brasileira de Pesquisa em educação em Ciências**, v. 5, n. 3, p. 38-51, 2005.

LEONARDO, André Ribeiro dos Santos. **Avaliação de desempenho:** Criação de um Sistema de Avaliação de Desempenho para uma empresa. 2013. Dissertação de Mestrado (Mestrado em Gestão Empresarial). Faculdade de Economia. Universidade do Algarve, Faro, Portugal, p. 114, 2013.

MONTE, Lucas Gonçalves. **Escala Likert difusa:** um estudo sobre diferentes abordagens. 2020. 44 f. Trabalho de Conclusão de Curso (graduação em Matemática Industrial) – Universidade Federal do Ceará, Centro de Ciências, Fortaleza, 2020.

MOREIRA, Marco Antônio. O que é afinal aprendizagem significativa? **Revista Currículum**, n. 25, p. 29-56, 2012.

MOREIRA, Marco Antônio; RIZZATTI, Ivanise Maria. Pesquisa em ensino. **Revista Internacional de Pesquisa em Didática das Ciências e Matemática**, [S. l.], v. 1, p. e020007, 2020.

PASQUALI, Luiz. Princípios de elaboração de escalas psicológicas. **Revista de Psiquiatria Clínica**, São Paulo, v. 25, n. 5, p. 206-213, 1998.

PUGLIESE, Renato Marcon. O trabalho do professor de Física no ensino médio: um retrato da realidade, da vontade e da necessidade nos âmbitos socioeconômico e metodológico. **Ciência & Educação (Bauru)**, v. 23, n. 1, p. 963-978, 2017.

REIS, Cleiton Pereira; MORALES, Juan Carlos Pérez. Validade de conteúdo da versão preliminar do instrumento para mensurar o conhecimento tático declarativo no basquetebol. **Revista Brasileira de Psicologia do Esporte**, v. 10, n. 2, p. 1-18, 2020.

RIBEIRO, Stefãne; MOREIRA, Alexandra; REIS, Janice; SOARES, Aleida; GÉa-HORTA, Tatiane. Elaboração e validação de cartilha sobre diabetes para Agentes Comunitários de Saúde. **Revista Brasileira de Enfermagem**, v. 73, n. 1, p. e20180899, 2020.

SANTOS, Flávia Maria Teixeira. As emoções nas interações e a aprendizagem significativa. **Ensaio Pesquisa em Educação em Ciências**, v. 9, n. 2, p. 173-187, 2007.

SILVEIRA, Joice Trindade; ROCHA, João Batista Teixeira. Tradução e ferramenta de validação de conteúdo de avaliação do ensino pelos alunos. **Revista Iberoamericana de Evaluación Educativa**, v. 10, n. 2, p. 45-64, 2017.

TAMADA, Rosane Cristina Piedade; CUNHA, Isabel Cristina Kowal Olm. Competências profissionais do técnico administrativo em educação: evidências de validade do conteúdo. **Revista Gestão Universitária na América Latina**, v. 16, n. 1, p. 01-22, 2023.

TROJAN, Rose Meri; SIPRAKI, Robson. Perspectivas de estudos comparados a partir da aplicação da escala Likert de 4 pontos: um estudo metodológico da pesquisa TALIS. **Revista Ibero-Americana de Estudos em Educação**, v. 10, n. 2, p. 275-300, 2015.

VALADARES, Jorge. A teoria da aprendizagem significativa como teoria construtivista. **Aprendizagem Significativa em Revista**, v. 1, n. 1, p. 36-57, 2011.

VIZZOTTO, Patrick Alves. **A proficiência científica de egressos do Ensino Médio ao utilizar a Física para interpretar o cotidiano do trânsito**. 2019. Tese (Doutorado em Educação em Ciências). Instituto de Ciências Básicas e da Saúde. Programa de Pós-graduação em Educação em Ciências. Universidade Federal do Rio Grande do Sul, Porto Alegre, p. 287, 2019.

YUSOFF, Muhamad Saiful Bahri. ABC of content validation and content validity index calculation. **Educational Resource**, v. 11, n. 2, p. 49-54, 2019.

Submissão: 28/11/2024

Aceito: 08/01/2025