

ANÁLISE ESTATÍSTICA DO DESEMPENHO DOS ITENS DE CIÊNCIAS DA NATUREZA NO ENEM: UMA COMPARAÇÃO DOS ANOS DE 2010 A 2020

Statistical analysis of the performance of nature science items on ENEM: a comparison between 2010 and 2020

Fernando Augusto Silva (fernando.augusto@ifmt.edu.br)

IFMT – Campus Avançado Guarantã do Norte

Linha Páscoa, Km 04, Lote 471, Zona Rural - CEP: 78520-000

Alexandre Antonello (aantonello22@gmail.com)

CPV Educacional

R. Domingos Lopes da Silva, 34 - Morumbi, São Paulo - SP, 05641-030

Recebido em: 12/08/2023

Aceito em: 29/11/2023

Resumo

O Exame Nacional do Ensino Médio (ENEM) é uma avaliação nacional que foi implementada no Brasil em 1998 com o objetivo inicial de avaliar a qualidade do ensino médio no país. No entanto, a partir de 2009, o ENEM também passou a ser utilizado como uma forma de acesso ao Ensino Superior, ampliando a importância e impacto no cenário educacional brasileiro. Diante desse contexto, muitos trabalhos têm discutido sobre o papel pedagógico e seletivo do Enem. Neste sentido, este trabalho se propôs a analisar de maneira conjunta, no que tange a Ciências da Natureza (Biologia, Física e Química), quais itens ao longo de dez anos (2010 a 2020), a partir do índice de acertos da Teoria Clássica dos Testes, são mais difíceis estatisticamente. Adicionalmente, este trabalho também busca relacionar a análise com as habilidades e competências descritas nas bases de dados do Inep. De modo geral, revelou-se que as disciplinas de Física e Química são as mais difíceis, pelo baixo índice de acertos e tornou-se cada vez mais seletiva e excludente ao longo dos anos. Portanto, é fundamental que os órgãos responsáveis pela elaboração do ENEM e outras avaliações nacionais estejam atentos à qualidade dos itens utilizados e à adequação em relação ao que é ensinado nas escolas. É preciso que haja um equilíbrio entre a seletividade e a função inclusiva da prova, garantindo que os estudantes concluintes do Ensino Médio tenham uma avaliação que reflita, de fato, o conhecimento adquirido durante o período escolar.

Palavras-chave: Ciências da Natureza; ENEM; Competências e Habilidades.

Abstract

The Exame Nacional do Ensino Médio (ENEM) is a national assessment that was implemented in Brazil in 1998 with the initial objective of evaluating the quality of secondary education in the country. However, as from 2009, the ENEM also started being used as a way of accessing Higher Education, increasing its importance and impact in the Brazilian educational scenario. Given this context, many works have discussed the pedagogical and selective role of the ENEM. In this sense, this work proposed to jointly analyze, regarding Natural Sciences (Chemistry, Physics and Biology), which items over ten years (2010 to 2020) are statistically more difficult. This work also seeks to relate to the skills and competencies described in the Inep databases. In general, it turned out that the subjects of Chemistry and Physics are the most difficult and that the test has become increasingly selective and exclusive over the years. Therefore, it is essential that the bodies responsible for preparing the ENEM and other national assessments pay attention to the quality of the items used and their adequacy in relation to what is taught in schools. There needs to be a balance between the selectivity and the inclusive function of the test, ensuring that students graduating from high school have an assessment that reflects their knowledge acquired during the school period.

Keywords: Natural Sciences; ENEM; Skills and Abilities.

Introdução

O Exame Nacional do Ensino Médio (ENEM) é uma avaliação nacional que foi implementada no Brasil em 1998 com o objetivo inicial de avaliar a qualidade do ensino médio no país. No entanto, a partir de 2009, o ENEM também passou a ser utilizado como uma forma de acesso ao Ensino Superior, ampliando sua importância e impacto no cenário educacional brasileiro. Com essa mudança, o exame ganhou maior relevância e tornou-se um indicador fundamental para estudantes que buscam ingressar em universidades e instituições de ensino superior (Silva, 2015). Além disso, o ENEM passou a ser considerado uma ferramenta que permite avaliar o desempenho dos estudantes em diferentes áreas do conhecimento, incluindo as Ciências da Natureza, que englobam disciplinas como Biologia, Física e Química.

Diante desse contexto, diversos pesquisadores têm se dedicado a estudar o ENEM em suas diversas perspectivas, investigando tanto sua estrutura e formato quanto suas possíveis correlações com currículos escolares, livros didáticos e outros aspectos educacionais (Peixoto, Linhares, 2010; Rubini *et al*, 2015; Viggiano, Mattos, 2015; Silva *et al*, 2017; Marcom, 2019; Silva, 2019; Bardini *et al*, 2021; Silva *et al*, 2022; Vizzoto, 2023). Essas pesquisas têm contribuído para a compreensão mais aprofundada do exame e para a busca de melhorias em sua aplicação e utilização. Nesse sentido, o presente estudo propõe uma análise estatística do desempenho dos itens de Ciências da Natureza no ENEM, focalizando a comparação dos anos de 2010 a 2020. Nosso objetivo é questionar qual área do conhecimento das Ciências da Natureza - Biologia, Física e Química - apresentou maior dificuldade ao longo dessa década.

Em vista de responder este questionamento, propomos realizar um breve resumo de alguns trabalhos que discutem sobre o ENEM, em suas perspectivas de análise. Posteriormente, também de maneira breve, iremos fazer uma discussão sobre a Teoria Clássica dos Testes (TCT) e suas contribuições para a análise de itens como ENEM. Seguidamente, para alcançar esse objetivo, utilizaremos técnicas estatísticas robustas e rigorosas, a fim de analisar e comparar os resultados obtidos nas questões relacionadas a essas disciplinas ao longo do período investigado. Essa análise proporcionará uma compreensão mais detalhada das áreas de conhecimento que podem representar um desafio maior para os estudantes que realizaram o ENEM. Por fim, a relevância desse estudo reside na contribuição para uma melhor compreensão do desempenho dos estudantes nas Ciências da Natureza, bem como no fornecimento de subsídios para a formulação de políticas educacionais mais efetivas, voltadas ao ensino dessas disciplinas e à preparação adequada dos alunos para o ENEM.

Fundamentação Teórica: Alguns trabalhos sobre Enem

Os autores Romão *et al* (2021) analisam 87 questões de Química do Enem de 2016 a 2019, em todas as suas aplicações, em relação a contextualização, pois, segundo eles, a contextualização desempenha um papel de grande importância, ao buscar aproximar os alunos dos conteúdos de forma concreta, tornando-os mais acessíveis, ou seja, essa conexão facilita o entendimento e a interpretação dos conceitos abordados. Em seus resultados, mostram que apenas 26%, em média, das provas aplicadas ao longo dos anos, são descontextualizadas e, ainda, que as questões que abordam os conceitos de Propriedade dos Materiais e Química Ambiental são as mais contextualizadas. Por fim, concluem que a contextualização está em consonância com a Lei de Diretrizes e Bases da Educação (LDB) e os Parâmetros Curriculares Nacionais (PCN+), uma vez que enfatiza a necessidade de desenvolver a criticidade dos alunos e o entendimento dos contextos que os cercam.

Concomitantemente, Silva e Souza (2018), realizaram uma análise das questões de Química do ENEM no período de 2009 a 2017, em relação ao conteúdo e a abordagem, com o objetivo de identificar as habilidades e competências mais recorrentes nessas avaliações e sua relação com a Matriz de Referência do ENEM¹. A partir dessa análise, constataram que as questões de Química do ENEM abordam temas sociais de maneira contextualizada, estabelecendo conexões entre os conceitos químicos e questões relevantes da sociedade. Segundo eles, essas conclusões têm grande relevância para os educadores, pois fornecem informações importantes para o planejamento de suas práticas de ensino. Eles reforçam que compreender as habilidades mais enfatizadas nas avaliações do ENEM permite que os professores preparem seus alunos de forma mais eficaz, promovendo uma base sólida de conhecimento químico e incentivando a reflexão crítica sobre os desafios sociais e ambientais presentes em nosso contexto.

Já os autores como Bardini, Brahm e Pastoriza (2021) problematizam o objetivo do ENEM como instrumento de avaliação, e realizam uma análise de 277 questões do Enem, de 1998 a 2019, a partir da metodologia do tetraedro de Sjöström (2013). Segundo os autores, os dados sugerem que o foco das questões de Química não está direcionado para avaliar um pensamento crítico que relacione os conceitos químicos com situações e fenômenos reais. Além disso, falta incentivo para que os alunos possam utilizar esse conhecimento químico em suas vivências e intervir em situações reais. Os autores destacam a relevância do nível 1, denominado pelos autores como Química Aplicada. Esse nível é essencialmente caracterizado como: “o ensino de conceitos e conteúdos através da exemplificação, em que o contexto citado somente é utilizado para uma aplicação direta dos conceitos” (pág. 3 e 4). Assim sendo, os autores revelam uma densidade conceitual não contextualizada, limitando-se a exemplos isolados. Diante disso, questiona-se qual é o direcionamento do exame com essa abordagem.

Na mesma direção, os autores Souza e Brito (2019), a partir do reconhecimento de que a abordagem educacional CTS (Ciência, Tecnologia e Sociedade) pode contribuir para o ensino de Química, pois estabelece conexões entre os conhecimentos escolares, o desenvolvimento científico e tecnológico, bem como sua aplicação na sociedade, visando a formação de cidadãos cientificamente alfabetizados, analisaram 156 questões relacionadas ao ensino de Química no novo Exame Nacional do Ensino Médio (ENEM). Dentro desse conjunto, constatou-se que 46,1% das questões não apresentaram afinidade com os princípios do enfoque CTS (Ciência, Tecnologia e Sociedade), enquanto 32,1% mostraram uma relação aparente com esses princípios e 21,8% estavam associadas aos mesmos. No estudo, foram identificadas três categorias de questões no ENEM, a saber: questões associadas aos pressupostos do enfoque educacional CTS; questões que apresentam dados científicos e tecnológicos, mas que não exigem interpretações qualitativas, nem abordam questões sociais; e questões que não apresentam dados científicos e tecnológicos, nem exigem interpretações qualitativas ou abordam questões sociais. Além disso, os resultados do estudo corroboram as considerações de outras pesquisas que apontam para a necessidade de uma abordagem mais problematizadora e reflexiva do ensino de Química no ENEM.

Os autores Miranda *et al* (2019) analisam as questões de Fisiologia Humana presentes no ENEM entre 1998 e 2016, primeiramente, utilizando a metodologia de Análise de Conteúdo de Bardin e, posteriormente, uma análise estatística. Através dessa primeira análise, o sistema imunitário foi o mais exigido nas questões de Fisiologia Humana ao longo das edições analisadas do ENEM. Os autores destacam que esse sistema foi mais exigido que outros sistemas igualmente importantes, como o sistema nervoso, cardiovascular e respiratório. Ademais, os autores também identificaram que as questões de Fisiologia Humana do ENEM abordaram principalmente processos fisiológicos e bioquímicos, em detrimento de processos evolutivos e de análise de impactos ambientais e que estas

¹ Disponível em: https://download.inep.gov.br/download/enem/matriz_referencia.pdf Acesso em 21 de nov.2023.

questões se concentram nas habilidades requeridas pelas competências 4 e 8 da Matriz de Referência da área de Ciências da Natureza. Ainda, destacam a importância da contextualização e da valorização da análise crítica em detrimento da memorização por repetição nas questões do ENEM.

Em termos de análise estatística, os autores encontraram que as questões de Fisiologia Humana apresentaram uma alta dispersão, com um Coeficiente de Variação (CV) de 74,50%, enquanto as questões de Biologia apresentaram uma média dispersão², com um CV de 21,76%. Tal fato sugere que o número de questões de Biologia do ENEM apresentou uma tendência mais homogênea que o número de questões de Fisiologia Humana. Por fim, eles sugeriram que futuros estudos possam investigar a relação entre as questões de Fisiologia Humana do ENEM e o desempenho dos estudantes nessa disciplina.

Por outro lado, Ferry, Vieira e Vidigal (2017) analisam 450 itens de Ciências da Natureza, nas edições de 2009 a 2016, a partir da estrutura dos itens, buscando identificar os tipos de modelos presentes na prova, assim, segundo os autores, existem os modelos didáticos e os modelos científicos. Observou-se, logo, que, em relação à subárea de conhecimento, a Biologia apresentou a menor quantidade de itens baseados em modelos em comparação com Física e Química. Logo, a análise específica para a área de Biologia indica uma menor utilização de modelos didáticos e científicos em questões de múltipla escolha, em comparação com as outras subáreas de conhecimento. Eles também destacaram a importância da utilização de modelos em avaliações em larga escala, como o ENEM, para desafiar os alunos a interpretar e/ou propor representações de diferentes entidades de interesse científico e, ao mesmo tempo, contribuir para o desenvolvimento crítico e flexível que pode ser aplicado em diferentes situações e problemas.

Assim, Vizzotto (2022) aborda a importância e a necessidade de garantir a qualidade das questões de Biologia no Exame Nacional do Ensino Médio (ENEM), pois é um exame de grande escala cuja nota é crucial para o ingresso em Instituições de Ensino Superior (IES) públicas e privadas, tendo um impacto significativo na vida profissional e social dos participantes. A qualidade das questões do ENEM é avaliada por meio de termos como "validade" e "fidedignidade", que fazem parte da área da psicometria. Nesse contexto, os parâmetros psicométricos são essenciais para verificar a qualidade dos itens das provas.

Neste sentido, a pesquisa de Vizzotto (2022) buscou analisar a qualidade psicométrica das questões de Biologia do ENEM nas edições de 2009 a 2019, cujo resultado indicou que a consistência interna das provas nas edições de 2013 e 2018 ficou abaixo do recomendado, apontando para a necessidade de melhorias nesse aspecto. Além disso, todas as provas foram classificadas como difíceis, o que pode indicar desafios enfrentados pelos participantes. Quanto à qualidade psicométrica, aproximadamente 54% das questões de Biologia não atenderam aos critérios esperados. Entende-se, portanto, que a pesquisa contribui para a literatura da área ao trazer evidências sobre a qualidade das questões de Biologia no ENEM, fornecendo informações relevantes para professores, pesquisadores e gestores educacionais.

Particularmente, no que tange ao ensino de Física, muitos são os trabalhos que discutem, em diversos aspectos, sobre o ENEM. Entretanto, de modo mais pragmático, o trabalho de Silva e Kawamura (2018), apresenta um panorama das pesquisas sobre o ENEM, que busca compreender o que é discutido no âmbito do ensino de Física. Desse modo, analisam 51 trabalhos publicados em

² Classifica-se como baixa dispersão de 0 a 15%, que indica baixa variabilidade relativa em relação à média. De 15% a 30%, classifica-se como média dispersão, que indica uma variabilidade moderada em relação à média e, por fim, acima de 30% é alta dispersão, que indica uma alta variabilidade em relação à média (Garcia, 1989)

eventos e em periódicos da área e propõem, a partir da metodologia de Análise de Conteúdo (Bardin, 2011), sistematizar em quatro categorias, a saber:

Categoria 1A – Análise da estrutura dos itens (Por exemplo, contextualização, interdisciplinaridade, objetos do conhecimento, competência, habilidade, conceitos abordados, tamanho do texto, mudanças após 2009). Categoria 1B – Análise nos itens, mas com relação ao desempenho dos estudantes (análises estatísticas, estudo dos distratores, etc.) Categoria 2 – Proposta do Enem: o foco não são os itens, mas a prova e/ou a sua proposta (comparação entre os conteúdos abordados e currículos, comparação com livro didático, comparação entre compreensões em trabalhos anteriores sobre a contextualização). Categoria 3 – Dimensões políticas: foco numa análise socioeconômica de desempenho, estrutura geral e qualidade dos itens, etc. (Silva, Kawamura, 2018, pág.5)

Ainda, constatam que, dos trabalhos analisados, 53% estão na categoria 1A. Em seguida, com apenas 20% os trabalhos da categoria 1B. Evidenciando, primeiramente, que há uma preocupação da área em discutir aspectos para além da estrutura da prova, como aspectos que envolvem dimensões políticas e dimensões pedagógicas. Além da necessidade de que haja mais trabalhos, principalmente nas categorias 1B, 2 e 3 (análise dos itens, proposta do Enem e dimensões políticas, respectivamente), sobre o ENEM.

Ademais, concatenado aos outros artigos, da área de Química e de Biologia, em que, sejam os que foram brevemente apresentados aqui ou são publicados em outros periódicos, em sua maioria analisam a estrutura dos itens e buscam dialogar com as diferentes práticas (como CTS) ou conteúdos (Souza, Pereira, 2015; Santos, *et. al.* 2016; Carvalho, *et. al.*, 2017; Freitas, Souza, 2018; Oliveira, *et. al.* 2019; Barbosa, *et. al.*, 2020; entre outros). Poucos são os trabalhos que dialogam em relação a estrutura do item, através de uma análise estatística e, também, são poucos os trabalhos que relacionam com as competências e habilidades da Matriz de Referência ENEM. Logo, todo este breve panorama reforça a necessidade de trabalhos que discutam aspectos estatísticos dos itens, como o que está sendo proposto, na busca de, entre outros aspectos, alimentar uma discussão maior do papel do ENEM para os respondentes de modo geral e, mais particularmente, para os concluintes do Ensino Médio.

Fundamentação teórica: Teoria Clássica dos Testes e Teoria da Resposta ao Item

Os parâmetros da Teoria Clássica dos Testes (TCT) são cruciais para garantir a qualidade dos itens em uma avaliação de larga escala. A TCT é uma abordagem estatística essencial para a avaliação educacional e pode fornecer *insights* valiosos sobre a eficácia dos itens de um teste. Para construir uma avaliação de larga escala que seja válida e confiável, alguns parâmetros podem ser considerados: o Índice de Dificuldade, o Ponto Bisserial e o Índice de Discriminação.

O Coeficiente ou Índice de Dificuldade é utilizado para mensurar o quão desafiador cada item era para os respondentes. Esse índice é calculado com base na proporção de respondentes que acertaram cada item em relação ao total de participantes da prova. Para que uma avaliação seja justa e abranja diferentes níveis de habilidades, é essencial incluir itens com diferentes níveis de dificuldade. Segundo Pasquali (2013) este índice também pode ser chamado de Índice de Facilidade, pois quanto maior a incidência sobre um item (mais próximo de 1), maior é a quantidade de respondentes que acertaram o item, de modo que se compreende como “itens fáceis”.

O Ponto Bisserial foi empregado como medida de correlação entre cada item e o desempenho geral dos estudantes na prova. Esse valor indica a capacidade do item em discriminar entre os participantes que obtiveram pontuações mais altas ou mais baixas. Ele varia de -1 a 1, onde valores mais próximos de -1 indicam uma relação negativa entre a resposta ao item e o *score* total do teste. Valores mais próximos de 1 indicam uma relação positiva e valores próximos de 0 indicam pouca

relação. Por outro lado, valores acima de 0,30 são geralmente considerados bons, de maneira tal que valores entre 0,20 e 0,30 indicam um item razoavelmente útil, mas que deve ser usado com cautela, e valores abaixo de 0,20 indicam um item com problemas de qualidade e que deve ser considerado para exclusão no teste (Hambleton, Swaminathan, 1985; Crocker, Algina, 1986; Anastasi, Urbina, 1997).

Por fim, o Índice de Discriminação, que avalia a capacidade do item em diferenciar os alunos com maior habilidade na disciplina em relação aos alunos com menor habilidade, é calculado a partir da correlação entre o desempenho no item e a pontuação total dos participantes. Desse modo, ele varia de -1 a 1, em que valores mais próximos de -1 indicam que o item é mais fácil para os participantes com habilidades mais baixas e mais difícil para os participantes com habilidades mais altas, valores mais próximos de 1 indicam o contrário, e valores próximos de 0 indicam que o item não é eficaz em diferenciar participantes com habilidades diferentes. Assim sendo, um valor positivo e significativo (geralmente acima de 0,40) indica que o item é útil para diferenciar participantes com habilidades diferentes na área medida (Hambleton, Swaminathan, 1985; Crocker, Algina, 1986; Anastasi, Urbina, 1997).

Por outro lado, a Teoria de Resposta ao Item (TRI) é uma abordagem estatística avançada frequentemente utilizada em avaliações educacionais, especialmente em testes de larga escala. Diferente da Teoria Clássica dos Testes (TCT), a TRI foca na modelagem das respostas individuais dos candidatos em relação às características dos itens do teste. Para garantir que uma avaliação seja válida e confiável, a TRI considera três parâmetros fundamentais para cada item do teste: o parâmetro de dificuldade (b), o parâmetro de discriminação (a) e o parâmetro de acerto ao acaso (c). (Pasquali, 2013)

Portanto, A TCT e a TRI são duas abordagens psicométricas complementares cujo objetivo é avaliar e melhorar a qualidade dos testes e itens utilizados em avaliações educacionais. Enquanto a TCT é uma abordagem mais tradicional e amplamente utilizada, a TRI é uma abordagem mais avançada que se baseia em modelos matemáticos para analisar as respostas individuais dos candidatos (Engel, 2018). Ambas as teorias possuem vantagens e limitações: a TCT é útil para verificar propriedades básicas dos itens, como a dificuldade, a discriminação e a consistência interna do teste e, portanto, o ponto de partida para uma análise básica dos itens; enquanto a TRI oferece uma vantagem significativa ao modelar as relações complexas entre as respostas dos candidatos e os itens, permitindo uma abordagem mais precisa e informativa sobre as habilidades dos candidatos e a qualidade dos itens.

Metodologia

A presente pesquisa utilizou uma abordagem metodológica quantitativa para realizar a análise estatística do desempenho dos itens de Ciências da Natureza no Exame Nacional do Ensino Médio (ENEM) nos anos pares, de 2010 a 2020, para fornecer uma amostra representativa e abranger um período significativo na investigação do desempenho dos itens. A coleta de dados foi realizada a partir do site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), garantindo a integridade e a confiabilidade das informações utilizadas na pesquisa. Utilizou-se, ainda, as informações disponíveis para a prova azul e os concluintes do Ensino Médio. Ao todo, foram analisados 268 itens relacionados às disciplinas de Biologia, Física e Química, conforme Tabela 1:

Tabela 1: Quantidade de itens para anos pares, de 2010 a 2020, para cada disciplina

Disciplina \ Anos	2010	2012	2014	2016	2018	2020	Total
Química	16	16	15	17	17	17	104
Física	14	15	15	15	13	14	81
Biologia	15	14	15	13	15	13	83
Total	45	45	45	45	45	44	268

O objetivo principal foi comparar os itens em relação a três medidas estatísticas de desempenho, a saber: o Índice de Dificuldade (Índice de Facilidade), o Ponto Bisserial e o Índice de Discriminação.

Para categorizar cada item, em relação ao Índice de Dificuldade, apoiou-se na classificação proposta por Gronlund (1974): percentuais de acerto de 80% a 100% caracterizam o item como muito fácil; de 60% a 79,9%, caracterizam o item como fácil; de 40% a 59,9%, o item é classificado como de média facilidade; de 20% a 39,9%, o item é considerado difícil; e de 0% a 19,9%, caracteriza o item como muito difícil.

Além da abordagem quantitativa, a pesquisa também adotou uma metodologia qualitativa-quantitativa (Marconi e Lakatos, 2017), para compreender, de forma interpretativa, a relação entre os níveis dos itens e as competências e habilidades contempladas em cada um deles. Essa análise permitiu uma compreensão mais aprofundada das características dos itens e sua relação com o currículo e os objetivos educacionais propostos. Dessa maneira, para realizar as análises estatísticas e a interpretação dos dados, foram utilizados softwares estatísticos e ferramentas específicas, garantindo a precisão e a confiabilidade dos resultados obtidos.

A abordagem qualitativa-quantitativa adotada na pesquisa permitiu uma análise mais ampla e completa dos resultados, pois além das medidas estatísticas, foi possível compreender a relação dos itens com as competências e habilidades esperadas dos estudantes. Essa análise combinada possibilitou uma interpretação mais aprofundada dos dados e uma compreensão mais contextualizada dos resultados obtidos.

Análise e discussão

A presente metodologia proporcionou uma análise abrangente e detalhada do desempenho dos itens de Ciências da Natureza no ENEM ao longo dos anos investigados, fornecendo informações valiosas para a compreensão das dificuldades enfrentadas pelos estudantes e das competências e habilidades demandadas nessas disciplinas. Neste sentido, iniciou-se a análise por uma abordagem mais geral e, em seguida, por disciplina. Posteriormente, foi realizada uma breve análise comparativa entre o desempenho dos itens e as competências e habilidades.

Um panorama geral

Inicia-se, apresentando no Gráfico 1, a classificação dos itens, a partir do desempenho dos concluintes do Ensino Médio, de acordo com o Índice de Dificuldade, segundo a TCT.

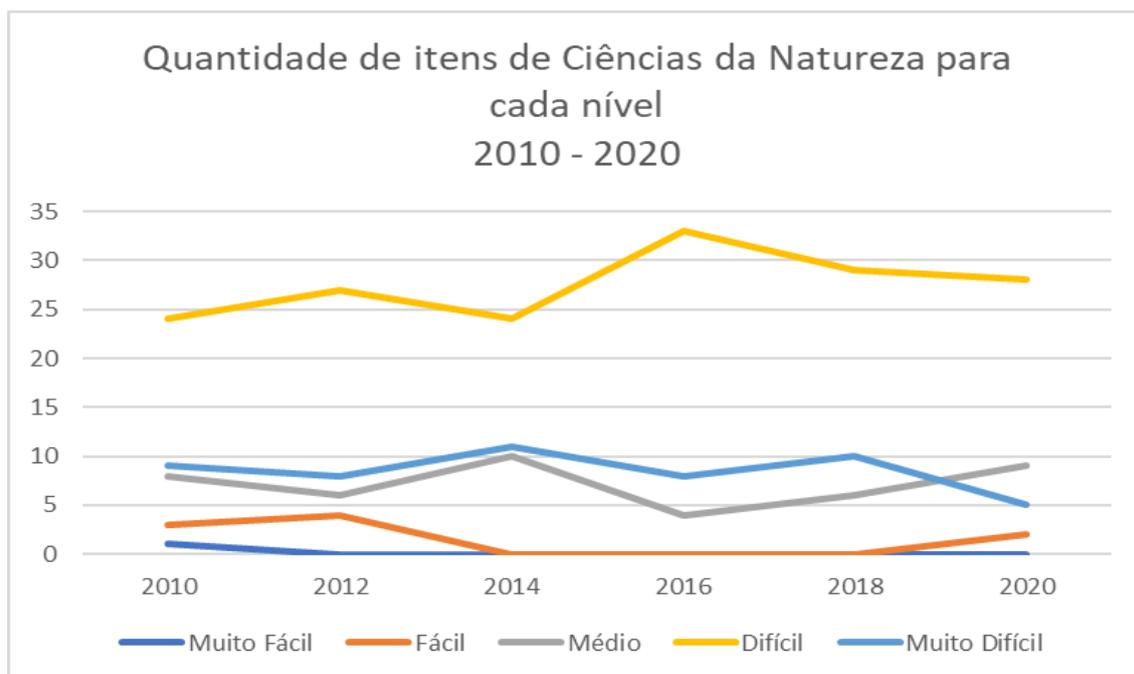


Gráfico 1: Classificação dos itens de Ciências da Natureza para cada nível, de 2010 a 2020, anos pares.

Observa-se, a partir do gráfico apresentado (Gráfico 1), que ao longo dos anos a quantidade de itens do nível difícil é consideravelmente maior, quando comparado as outras classificações, assim como a quantidade de itens fáceis e muito fáceis praticamente são nulas ao longo dos anos. Ainda, a quantidade de itens muito difíceis é maior do que a quantidade de itens médios, exceto em 2020.

Segundo Pasquali (2013):

“Pode-se, finalmente, perguntar ainda se existe um nível ideal de dificuldade para os itens de uma escala o teste. A resposta a esta indagação depende da finalidade do teste. (...)”

Se, entretanto, o interesse consiste em avaliar a magnitude diferencial dos traços nos sujeitos de uma população, como geralmente é o caso, então uma distribuição mais equilibrada dos itens em termos de dificuldades é requerida.” (pag. 127 e 128).

Assim, de modo geral, esta forma de construir um teste não coaduna com a proposta de uma prova que, por ser de escala abrangente, deveria haver mais itens da classificação média.

Uma análise por disciplina

Ao realizar a análise dos dados, foi possível identificar que os itens de Química e Física apresentaram os maiores níveis de dificuldade, com taxas de acerto de 22% e 25%, respectivamente, enquadrando-os na categoria de itens de nível muito difícil. Por outro lado, todos os itens demonstraram uma distribuição quase uniforme na categoria de dificuldade difícil, com aproximadamente 60% de acertos, conforme Tabela 2.

Tabela 2: Quantidade de itens, por anos pares, de 2010 a 2020, para cada disciplina, para os níveis muito difícil e difícil.

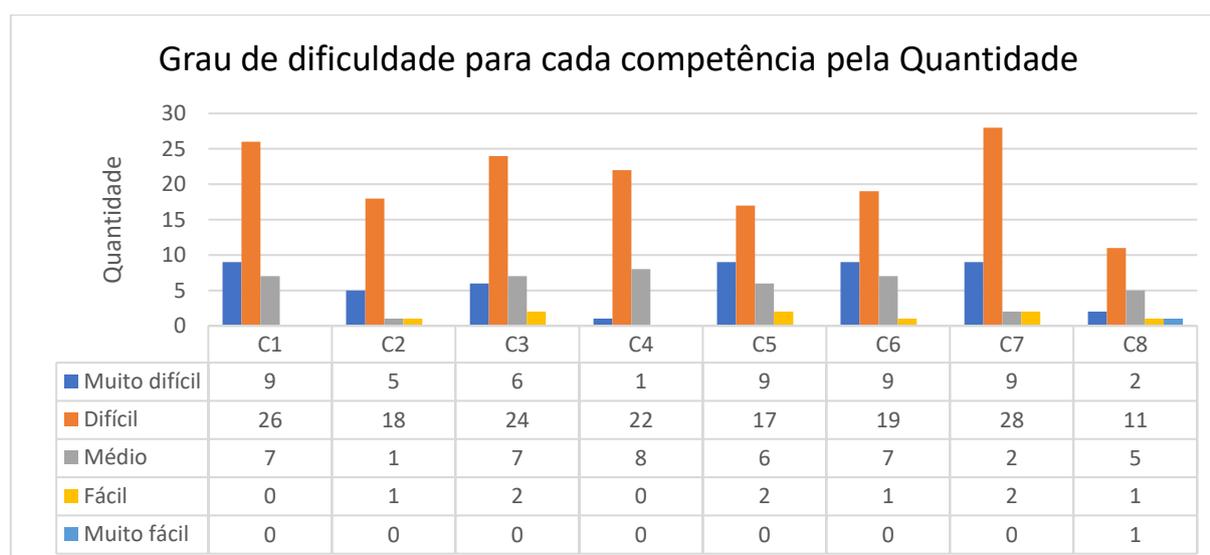
Nível	Disciplina	2010	2012	2014	2016	2018	2020	Total	%
Muito Difícil	Química	5	2	4	4	4	4	23	22%
	Física	3	5	6	4	3	1	22	27%
	Biologia	1	1	1	0	3	0	6	7%
Difícil	Química	10	11	9	11	12	10	63	61%
	Física	6	7	7	11	8	10	49	60%
	Biologia	8	9	8	11	9	8	53	64%

Em relação ao Ponto Bisserial, observou-se que 12 itens muito difíceis e 14 itens de dificuldade difícil apresentaram problemas. Isso sugere que esses itens podem não estar discriminando adequadamente entre os estudantes com pontuações mais altas e mais baixas, indicando a necessidade de revisão e possíveis ajustes nesses itens para garantir a sua qualidade e eficácia.

Finalmente, constatou-se que 24 itens apresentaram um valor de índice de discriminação inferior a 0,10. Esse resultado sugere que esses itens podem não ser eficientes em diferenciar os estudantes com maior habilidade na disciplina daqueles que possuem menor habilidade. Dessa forma, é recomendado que esses itens sejam revisados e, possivelmente, eliminados para garantir uma avaliação mais precisa e confiável.

Uma análise comparativa com as competências e habilidades

Por fim, ainda que de modo panorâmico, analisa-se a relação entre as competências contempladas no ENEM ao longo dos anos e o nível de dificuldade, conforme Gráfico 2.

**Gráfico 2:** Quantidade de itens, para cada nível de dificuldade, por competências da Matriz de Referência do Enem

Ao examinarmos as competências associadas aos itens classificados como muito difíceis, destacaram-se as competências 1, 5, 6 e 7³. Isso indica que essas competências são particularmente desafiadoras para os estudantes, demandando maior atenção no processo de preparação para o ENEM. Já para os itens de dificuldade difícil, as competências 1 e 7 foram as mais enfatizadas, sugerindo a necessidade de um foco especial nessas áreas durante a preparação.

Por fim, no que se refere às habilidades, podemos observar que - coerente a quantidade de itens -, em sua maioria se concentram no Índice de Dificuldade Difícil, conforme Tabela 3.

Tabela 3: Quantidade de habilidades para cada Índice de Dificuldade e a sua porcentagem.

ID	Quant de Hab	%
Muito Fácil	1	0%
Fácil	8	4%
Médio	34	15%
Difícil	139	62%
Muito Difícil	43	19%
Total	225	

Assim, identificou-se que as habilidades 1, 3, 6, 19, 20 e 22⁴ foram as mais destacadas nos itens classificados como muito difíceis. Essas habilidades específicas requerem um conhecimento mais aprofundado e uma maior capacidade de aplicação por parte dos estudantes. Por outro lado, para os itens de dificuldade difícil, as habilidades específicas de Química (24, 25 e 26) foram as mais comuns, enfatizando a importância do domínio dessas habilidades para um desempenho satisfatório.

Conclusão

Essa análise detalhada dos itens de Ciências da Natureza no ENEM ao longo dos anos investigados proporciona *insights* valiosos sobre as áreas de maior dificuldade e os aspectos que necessitam de atenção especial. Com base nesses resultados, é possível direcionar esforços para o desenvolvimento de estratégias de ensino mais eficazes, o aprimoramento dos materiais didáticos e a formulação de políticas educacionais voltadas para o fortalecimento dessas habilidades e competências específicas.

Desse modo, a análise dos itens de Ciências da Natureza no ENEM de 2010 a 2020 revelou que as disciplinas de Química e Física são as mais difíceis e que a prova se tornou cada vez mais seletiva e excludente ao longo dos anos. Logo, é fundamental que os órgãos responsáveis pela elaboração do ENEM e outras avaliações nacionais estejam atentos à qualidade dos itens utilizados e à sua adequação em relação ao que é ensinado nas escolas. É preciso que haja um equilíbrio entre a seletividade e a função inclusiva da prova, garantindo que os estudantes concluintes do Ensino Médio tenham uma avaliação que reflita de fato o seu conhecimento adquirido durante o período escolar. Além disso, é importante que sejam realizadas pesquisas contínuas sobre o desempenho dos estudantes nas diferentes áreas do conhecimento, para que possam ser implementadas políticas educacionais mais efetivas que vislumbrem uma educação de qualidade para todos.

³ Para saber quais são estas competências de Ciências da Natureza, indica-se consultar a Matriz de Referência Enem. Disponível em: https://download.inep.gov.br/download/enem/matriz_referencia.pdf (pág. 8 a 10). Acesso em 21 de nov. de 2023.

⁴ Para saber quais são estas habilidades de Ciências da Natureza, indica-se consultar a Matriz de Referência Enem. Disponível em: https://download.inep.gov.br/download/enem/matriz_referencia.pdf (pág. 8 a 10). Acesso em 21 de nov. de 2023.

É importante destacar, ainda, que a presente pesquisa tem como base a análise estatística dos dados disponíveis e, portanto, suas conclusões estão limitadas aos itens analisados e ao período investigado. Limitações adicionais podem incluir a falta de informações sobre outros aspectos relevantes do ENEM, como a influência de fatores socioeconômicos e educacionais no desempenho dos estudantes.

Por fim, ressalta-se que a utilização de uma metodologia rigorosa e detalhada contribui para a confiabilidade e a validade dos resultados obtidos nesta pesquisa, fornecendo subsídios para a compreensão do desempenho dos itens de Ciências da Natureza no ENEM ao longo dos anos analisados. Essa análise aprofundada pode ser útil para aprimorar as estratégias de ensino, o desenvolvimento de materiais didáticos e a elaboração de políticas educacionais mais eficazes nesta área do conhecimento.

Referências

- ANASTASI, A., & URBINA, S. **Psychological Testing**. 7. ed. Prentice Hall, 1997.
- BARBOSA, A. L.; SILVA, J. L.; SANTOS, M. G. Análise do desempenho nas questões de Química no ENEM: uma comparação entre os anos de 2010 a 2020. **Revista Brasileira de Pesquisa em Educação em Ciências**, v. 20, n. 3, p. 563-584, 2020.
- BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2011.
- BARDINI, L. S., BRAHM, G. S., PASTORIZA, B. S. Ensino de Química e ENEM: problematizando os níveis do conhecimento químico. **XIII Encontro Nacional de Pesquisa em Educação em Ciências – XIII ENPEC, ENPEC EM REDES**, 2021.
- CARVALHO, C. C.; FERREIRA, D. B.; RODRIGUES, E. C. Análise estatística do desempenho dos estudantes no ENEM: um estudo comparativo entre as áreas do conhecimento. **Revista Brasileira de Estudos Pedagógicos**, v. 98, n. 247, p. 550-569, 2017.
- CROCKER, L., & ALGINA, J. **Introduction to Classical and Modern Test Theory**. Holt, Rinehart and Winston, 1986.
- ENGEL, L. C. **Avaliação Educacional: Estatística Básica e Teoria da Resposta ao Item**. Editora SBS, 2018.
- FREITAS, R. L.; SOUZA, V. C. Desempenho dos estudantes nas áreas de Ciências da Natureza do ENEM: uma análise comparativa entre as disciplinas. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 26, n. 101, p. 369-390, 2018.
- GARCIA, C.H. **Tabelas para classificação do coeficiente de variação**. Piracicaba: IPEF, 1989. 12p. (Circular técnica, 171).
- GRONLUND, N. E. **A elaboração de testes de aproveitamento escolar**. São Paulo: EPU, 1974.
- HAMBLETON, R. K., & SWAMINATHAN, H. **Item Response Theory: Principles and Applications**. Kluwer-Nijhoff Publishing, 1985.
- MARCOM, Guilherme Stecca. **O ENEM, indicadores formativos e o ensino de física**. Tese (doutorado). Universidade Estadual de Campinas, Instituto de Física, Campinas, SP, 2019.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de Metodologia Científica**. 8. ed. São Paulo: Atlas, 2017.

MIRANDA, L. A. S. de; FERREIRA, A. C. F.; DIAS, G. R. M. Análise de conteúdo das questões de Fisiologia Humana da Prova de Ciências da Natureza e suas Tecnologias do Exame Nacional do Ensino Médio (1998-2016). **Ciência & Educação (Bauru)**, v. 25, n. 2, p. 375-393, 2019.

OLIVEIRA, M. C.; NASCIMENTO, F. S.; SANTOS, L. G. Análise do desempenho nas questões de Ciências da Natureza no ENEM: um estudo longitudinal de 2010 a 2018. **Educação & Realidade**, v. 44, n. 4, e104370, 2019.

PASQUALI, L. **Psicometria: teoria dos testes na psicologia e na educação**. 5. ed. Petrópolis: Editora Vozes, 2013.

PEIXOTO, K.C.Q.C, LINHARES, M.P. (2010) Novo ENEM: o que mudou? Uma investigação dos conceitos de Física abordados no Exame. **XII EPEF**, Águas de Lindóia.

ROMÃO, I. C. *et al*, Análise da contextualização dos conteúdos de Química nas questões do exame nacional do ensino médio (ENEM). **Brazilian Journal of Development**, (Curitiba), v.7, n.1, p. 10215-10224, jan 2021.

RUBINI, G., MASSUNAGA, M.S.O., BARROSO, M.F. É possível avaliar o que os alunos aprendem em Física usando o ENEM? **Anais do XXI SNEF**. São Paulo: SBF, 2015.

SANTOS, P. A.; COSTA, F. B.; LIMA, C. S. Avaliação do desempenho dos itens de Biologia no ENEM: uma análise comparativa entre os anos de 2010 e 2015. **Revista Electrónica de Enseñanza de las Ciencias**, v. 15, n. 3, p. 715-730, 2016.

SILVA, F.A. **As avaliações externas e as possibilidades de diálogo com os professores de Física**. Tese (doutorado). Programa de Pós-Graduação Interunidades em Ensino de Ciências, Universidade de São Paulo, 2019.

SILVA, F. A., KAWAMURA, M. R. D. Reflexões sobre o ENEM: contribuições da pesquisa em ensino de física. 2018, **Anais..** São Paulo: SBF, 2018. Disponível em: <https://sec.sbfisica.org.br/eventos/epef/xvii/sys/resumos/T0291-1.pdf>. Acesso em: 21 nov. 2023.

SILVA, J. M.; FIDELIS, K. R.; DA SILVA, M. A.; LIMA, K. E. C. A Biologia nas questões do Novo Enem: um estudo de alinhamento entre a Matriz de Referência e as edições de 2014 a 2019. **Revista Brasileira do Ensino Médio**, v. 5, p. 1-17, 2022.

SILVA, M. E. O ENEM e o acesso à educação superior no Brasil. **Revista do Instituto de Estudos Brasileiros**, v. 61, p. 203-221, 2015.

SILVA, R.C.F., SOUZA, V.C.A. Investigação das habilidades e competências trazidas nas questões de química do Enem 2009-2017 a partir da análise de conteúdo. **Revista Ciências & Ideias**, ISSN 2176-1477, vol. 9, n.3, 2018.

SILVA, V.A., MARTINS, M.I., SANTANA, M.T.S. Objetos de conhecimento de Física nas Edições 2009-2015 do Exame Nacional do Ensino Médio (ENEM). **Anais do XXII SNEF**. São Paulo, 2017.

SOUZA, J. R da T; BRITO, L, P de. Análise de aproximações das concepções presentes em questões associadas ao ensino de Química aplicadas no novo Enem com as concepções do enfoque CTS. **Revista de Educação em Ciências e Matemática, Amazônia**, v.15, n. 33, p.116-133, 2019.

SOUZA, L. M.; PEREIRA, L. A. Desempenho dos estudantes nas questões de Física no ENEM: uma análise comparativa entre os anos de 2010 e 2014. **Investigações em Ensino de Ciências**, v. 20, n. 3, p. 145-160, 2015.

VIGGIANO, E., MATTOS, C. A influência de características socioeconômicas no desempenho de participantes na prova de Ciências da Natureza do Enem 2012. **Anais do X ENPEC**. Rio de Janeiro: ABRAPEC, 2015.

VIZZOTTO, P. Análise e classificação das questões de Biologia do ENEM segundo suas características psicométricas. **Revista de Ensino de Biologia da SBEnBio**, [S. l.], v. 15, n. 1, p. 314–332, 2022. DOI: 10.46667/renbio.v15i1.633. Disponível em: <https://renbio.org.br/index.php/sbenbio/article/view/633>. Acesso em: 24 jul. 2023.