

ANALISANDO O IMPACTO DAS REPRESENTAÇÕES DE GÊNERO NO ENEM 2018

ANALYSING THE IMPACT OF GENDER REPRESENTATIONS IN THE ENEM 2018

Viviana da Cruz Vicente¹, André Machado Rodrigues²

¹ Universidade de São Paulo /Doutoranda no Programa de Pós-graduação Interunidades em Ensino de Ciências/Instituto de Física, vivianavic@usp.br

²Universidade de São Paulo/Professor Doutor no Programa de Pós-graduação Interunidades em Ensino de Ciências/Instituto de Física, rodrigues@if.usp.br

Resumo

A pesquisa se consistiu em uma investigação empírica realizada a partir dos dados e itens do Exame Nacional do Ensino Médio (ENEM) de 2018. Por meio da aplicação de algumas questões de Ciências da Natureza do exame, em uma escola da rede estadual paulista, foi possível identificar a percepção de estudantes quanto ao contexto dos enunciados apresentados. A partir dos microdados disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), e com utilização do programa R, foi gerado o Funcionamento Diferencial dos Itens (DIF) e analisadas as curvas características. Os resultados indicaram que um item sobre força, quando apresentado a homens e mulheres, foi avaliado como tendo representações de gênero masculinas e o grau de familiaridade que se sobressaía foi o médio. Em outro item, onde se discutia a lavagem de jeans, o contexto foi classificado como pertencente do universo feminino, média familiaridade para as mulheres e baixa para os homens. As questões também foram analisadas quanto a presença e ausência de DIF significativa. Apenas no primeiro, de força, que foi detectada a DIF.

Palavras-chave: Ensino de Física, Avaliação, ENEM, Gênero.

Abstract

The present search consists in an empirical investigation, also statistic, accomplished from information's about the Exame Nacional do Ensino Médio (ENEM) 2018. Thought an application of some Natural Science questions, in the public school in the state of São Paulo, it was possible to identify the perceptions of participants as to the context of the statements presented. From the microdata available on the Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) and using the statistic software R, were generated the Differential Item Functioning (DIF) and the analysed the characteristics curves. The results indicated that in one item about force, when presented for women and man, it was evaluated as having gender man representations and the degree of familiarity that stood out was the average. In other item, where jeans washing were discussed, the context was classified as belonging to the female universe and of average familiarity for women and low for men. The questions also were analyzed as for the presence and absence of DIF significant. Only the first, of the strength, that DIF was detected.

Keywords: Teaching Physics, Evaluation, ENEM, Gender.

Introdução

Essa pesquisa, de natureza empírica, buscou verificar se as representações de gênero, presentes em algumas questões do bloco de Ciências da Natureza do Exame Nacional do Ensino Médio (ENEM) de 2018, 1ª aplicação, eram próximas da realidade dos (as) estudantes. Outro aspecto analisado, nesse estudo, foi se a probabilidade de acertos em um item, por gênero, para indivíduos com a mesma proficiência, eram semelhantes ou se apresentavam diferenças. Uma das hipóteses é que as ilustrações e textos contidos nos enunciados podem favorecer um ou outro gênero. Desta forma, para identificar se há diferenças na probabilidade de acertos, investigou-se o Funcionamento Diferencial do Item. Em inglês, o termo é denominado Differential Item Functioning (DIF).

As análises da DIF geralmente envolvem dois campos: o pedagógico e o estatístico. O primeiro visa investigar se os itens elaborados estão de acordo com a habilidade que pretende medir e, o segundo, se eles possuem os parâmetros necessários para um bom desempenho no instrumento de medição (FACHINETTO, 2016).

Bateson e Parsons-Chatman (1989) afirmam que as desigualdades observadas no item se devem, em partes, as características deles medirem as diferenças dos estudantes fora das experiências da escola e, também, o conhecimento nas outras áreas. Deste modo, visando produzir uma prova que possibilite com que os itens tenham o mesmo nível de dificuldade para ambos os gêneros, são bem vindas as pesquisas que examinem-os individualmente. Buscando, assim, identificar quais aspectos contribuem para reduzir as desigualdades de gênero nos testes padronizados.

Sudweeks e Tolman (1993), mencionam que os recursos utilizados para identificar excertos potencialmente tendenciosos incluem métodos de julgamento e estatística. O primeiro, é subjetivo e visa examinar os itens com a finalidade de reconhecer estereótipos, cultura específica ou ofensiva objetivando eliminá-la. A pretensão é garantir com que os grupos sejam representados de maneira justa e que os enunciados retratem os estilos de vida dos grupos.

Considerando que as questões do ENEM são contextualizadas, e visam identificar as habilidades, o DIF possibilita com que o item seja investigado em relação ao nível de dificuldade e qualidade. Assim, quando os itens apresentam DIF significativa, pode-se dizer que há desigualdades de gênero na probabilidade de acertos e, portanto, que o instrumento mede de maneiras distintas a aptidão de cada grupo.

1. O ENEM

O ENEM, criado em 1998, é realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Baseando-se em testes de proficiência, a prova visa medir o desempenho discente quanto as habilidades consideradas fundamentais em cada disciplina. Inicialmente, até 2008, a finalidade do exame era a de avaliar o desempenho escolar na etapa final da escola básica e fornecer elementos para a elaboração de políticas educacionais.

Em 2005, com a criação do Programa Universidade para todos (PROUNI), bolsas começaram a ser ofertadas conforme a nota do ENEM e a classificação obtida no processo. No segundo formato, a partir de 2009, Marcom e Kleinke (2017, p. 46) mencionam que o exame “tornou-se também instrumento avaliativo de acesso ao ensino superior - semelhante ao SAT nos Estados Unidos da América”. No mesmo ano, a prova que anteriormente tinha sessenta e três itens, passou para cento e oitenta e uma redação. A partir desse ano, a nota no exame se tornou requisito obrigatório para o ingresso nas Instituições Federais de Ensino. O Sistema de Seleção Unificada (SiSU), para as pessoas que almejavam uma vaga nas instituições, classificava participantes conforme a performance.

De acordo com a literatura, o exame é produzido a partir de Matrizes de referência (FERREIRA, 2018, p. 20). Estas são constituídas por competências de quatro áreas de conhecimento: “Linguagens, Códigos e suas tecnologias, Matemática e suas tecnologias, Ciências Humanas e suas tecnologias e Ciências da Natureza e suas tecnologias, apresentando ainda, eixos cognitivos comuns a todas as áreas”.

Blasis, Falsarella e Alavarse (2013) mencionam que as avaliações mostram informações sobre os resultados obtidos pelas escolas e redes de ensino. Baseando-se no desempenho nas provas padronizadas, é possível identificar o que foi aprendido, ou não, e fazer inferências sobre o trabalho educativo realizado dentro do contexto escolar.

2. Metodologia

2.1 Caracterização do público alvo e coleta de dados

Em relação a escola participante, a presente possui os seguintes níveis de ensino: Fundamental II, Médio (regular) e modalidade EJA (Educação para Jovens e Adultos). O público alvo do trabalho, constituído apenas por discentes do nível médio (EJA e regular), eram pessoas entre 14 e 61 anos. No total foram cento e oitenta e seis pessoas que fizeram parte do estudo. No que se refere a escolha dos (as) estudantes que preencheriam os formulários, essa pesquisa não se preocupou em delimitar os gêneros ou idades.

A pesquisa foi produzida em duas partes constituídas pela aplicação de formulários e pela interpretação de gráficos gerados a partir dos microdados do ENEM. Na primeira etapa, discentes receberam alguns impressos contendo oito questões do bloco de Ciências da Natureza e deveriam responder duas perguntas: se o enunciado abordava mais elementos presentes no universo feminino, masculino, indeterminado ou não sabia e o quanto o contexto do mesmo era familiar para ela (em uma escala de 1-Muito até 5-Pouco). O critério de seleção, destas questões, foi a existência de objetos facilmente encontrados no cotidiano e que de algum modo pareciam se aproximar mais de um dos universos masculino, feminino ou de ambos. Adicionalmente a isto, as ações presentes nos enunciados também foram levadas em consideração.

Das oito questões, após a análise das respostas dadas aos instrumentos que se baseou na escala Likert (onde são atribuídos pontos para cada categoria), as de número 104 e 113 do caderno azul (1ª aplicação de 2018) foram selecionadas para serem geradas as curvas características dos itens. A segunda etapa se consistiu na investigação dos microdados disponibilizados no sítio do INEP,

possibilitada a partir do programa R, e na interpretação da curva gerada. Buscando, assim, identificar o *Differential Item Functioning* (DIF).

2.2 Teoria de Resposta ao Item (TRI) e o *Differential Item Functioning* (DIF)

A Teoria de Resposta ao Item (TRI), no ENEM, é utilizada pelo INEP desde 2009. Por meio dela, torna-se possível medir o nível de proficiência da pessoa participante por área de conhecimento. Sendo considerado 500 o desempenho médio e 100 a medida de variabilidade média das notas (BRASIL, 2015). A partir da representação gráfica, que expressa a proficiência e também a característica por item, é possível interpretar os resultados pedagogicamente.

De acordo com Carlos (2016, p. 38) existem diversos modelos matemáticos que permitem expressar a probabilidade de cada participante “acertar ao item em função da proficiência”. Entre esses, segundo Andrade, Tavares e Valle (2000), o mais utilizado pela TRI é o modelo unidimensional logístico de 3 parâmetros (ML3). De acordo com os autores descritos, matematicamente, podemos representá-lo pela seguinte expressão:

$$P_{ij} = \frac{1}{1 + e^{-(\theta_j - \beta_i - D)}} \quad (1)$$

Sendo que: $i = 1, 2, \dots, I$, e $j = 1, 2, \dots, n$; j é a representação de um indivíduo; i representa um item; P_{ij} é uma variável dicotômica que assume os valores 1, quando j acerta o item i , ou 0 quando j erra; θ_j é a habilidade do j -ésimo indivíduo; P_{ij} é a probabilidade de j com habilidade θ_j acertar o item i ; β_i é o parâmetro de dificuldade; α_i é o parâmetro de discriminação; c_i é o parâmetro de resposta ao acaso (possíveis chutes); D é um fator de escala, constante e igual a 1.

No modelo acima (ML3), a relação entre o indivíduo acertar i e os parâmetros desse item, é uma função de ordem crescente conhecida por Curva Característica do Item (CCI). Considerando que o Funcionamento Diferencial do Item, DIF, é percebido por meio dessa curva, julga-se necessária a descrição dos elementos da TRI no trabalho.

Os esforços em compreendê-la, originados com a psicometria moderna, visavam contribuir para a produção de testes imparciais. Cole (1993), defende que a preocupação com o DIF está relacionada com a vontade de construir questões que não fossem impactadas pelas características culturais dos grupos participantes dos testes que fariam parte de uma avaliação educacional.

De acordo com Fachinetti (2016, p.66), a análise da DIF é apoiada em estudos comparativos realizados após o estabelecimento de grupos. Na maioria das pesquisas, associadas à temática, existem duas populações: um Grupo de Referência (GR) e um Grupo Focal (GF). Esse primeiro (GR), geralmente é o preponderante.

No que se refere ao domínio da TRI, Aguiar (2010, p. 173) menciona que: “é possível dizer que o item não tem DIF, quando a curva (CCI) é a mesma para os grupos comparados em um mesmo nível de habilidade ou proficiência medida pelo item”. Caso contrário, se a curva (CCI) for diferente entre tais grupos, o item apresenta DIF.

Mellemborg (1982) menciona que a DIF pode ser classificada como uniforme e não uniforme. Em acordo com a literatura da área, segundo Andriola (2000), Mellemborg (1982) e Valle (2002), ela é uniforme quando as CCI dos grupos estudados (GR e GF) são diferentes e não se cruzam. À respeito da DIF não uniforme, o autor descreve que é quando as CCI dos grupos estudados são diferentes e se cruzam em algum ponto do intervalo. Além dessa categorização, Douglas, Roussos e Stout (1996), propuseram duas classificações associadas a DIF. Segundo os autores, quando a DIF beneficia o GR é denominada de Benigna e no caso de favorecer o GF ela é chamada de Adversa.

3. Análise de dados e resultados

As imagens 1 e 2 se referem as questões 104 e 113 que integraram o formulário aplicado. Na primeira, que trouxe à tona elementos como alicate, pinça, carrinho de mão, quebra-nozes e abridor de garrafa, a representação de gênero identificada no enunciado, e que obteve maior destaque, foi a masculina. No que diz respeito ao grau, tanto para homens quanto para mulheres, a proximidade com a situação foi classificada como média. Na segunda, sobre clareamento e lavagem de jeans, a opção que se destacou foi a feminina. Quando a familiaridade do respondente, em relação a descrição dada na questão, a classificação obtida foi de média familiaridade para o público feminino e pouca para o masculino.

QUESTÃO 104

As pessoas que utilizam objetos cujo princípio de funcionamento é o mesmo do das alavancas aplicam uma força, chamada de força potente, em um dado ponto da barra, para superar ou equilibrar uma segunda força, chamada de resistente, em outro ponto da barra. Por causa das diferentes distâncias entre os pontos de aplicação das forças, potente e resistente, os seus efeitos também são diferentes. A figura mostra alguns exemplos desses objetos.



Em qual dos objetos a força potente é maior que a força resistente?

- ☐ A Pinça.
- ☐ B Alicate.
- ☐ C Quebra-nozes.
- ☐ D Carrinho de mão.
- ☐ E Abridor de garrafa.

Imagem 1-Questão 104 da 1ª aplicação do caderno azul 2018-2º dia de prova

QUESTÃO 113

Companhias que fabricam jeans usam cloro para o clareamento, seguido de lavagem. Algumas estão substituindo o cloro por substâncias ambientalmente mais seguras como peróxidos, que podem ser degradados por enzimas chamadas peroxidases. Pensando nisso, pesquisadores inseriram genes codificadores de peroxidases em leveduras cultivadas nas condições de clareamento e lavagem dos jeans e selecionaram as sobreviventes para produção dessas enzimas.

TORTORA, G. J.; FUNKE, B. R.; CASE, C. L. Microbiologia. Rio de Janeiro: Artmed, 2016 (adaptado).

Nesse caso, o uso dessas leveduras modificadas objetiva

- ☐ A reduzir a quantidade de resíduos tóxicos nos efluentes da lavagem.
- ☐ B eliminar a necessidade de tratamento da água consumida.
- ☐ C elevar a capacidade de clareamento dos jeans.
- ☐ D aumentar a resistência do jeans a peróxidos.
- ☐ E associar ação bactericida ao clareamento.

Imagem 2-Questão 113 da 1ª aplicação do caderno azul 2018-2º dia de prova

A partir de um levantamento do DIF para os itens 104 (gráfico 3) e 113 (gráfico 4), as curvas para os itens citados foram geradas.

O eixo y, dos gráficos 3 e 4, representa a probabilidade de resposta correta (), ou seja, de zero até cem por cento (convertida na escala de 0,00 até 1,00). O eixo x, denominado escore padronizado, exibe a habilidade (), . A linha contínua, em azul, é o grupo de referência (GR) e corresponde as mulheres. A linha tracejada, em laranja, é o grupo focal (GF) e se refere as respostas dadas pelos homens.

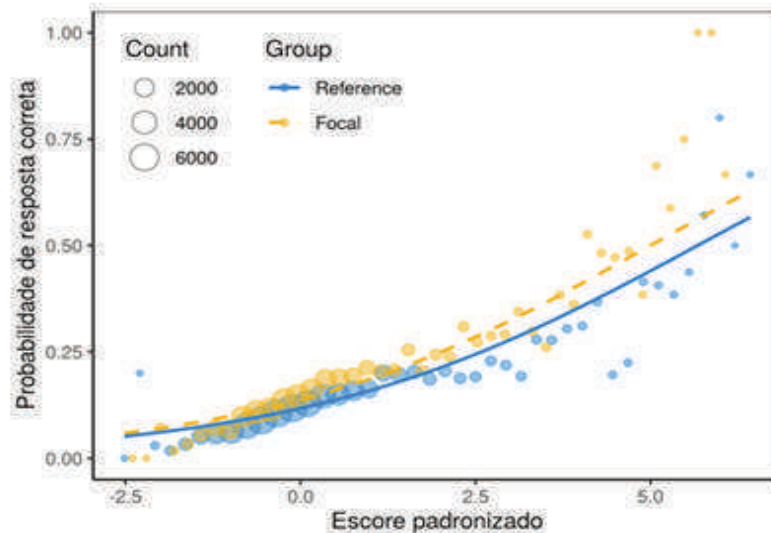


Gráfico 3-Questão 104 da 1ª aplicação do caderno azul 2018-2º dia de prova

No gráfico 3, a CCI dos homens (GF) está localizada mais a esquerda que a das mulheres (GR), indicando que o item (aproximadamente a partir do escore padronizado 2.0) parece ser mais fácil para o grupo focal. Essa diferença no parâmetro b, de dificuldade, permite supor que o item possui DIF significativa e que pode ser inferida a partir da distância vertical do GR e GF. Neste caso, o item provavelmente é considerado mais difícil pelo grupo de referência (composto pelas mulheres). Para o mesmo gráfico, como as curvas não se cruzam, pode-se afirmar que a DIF é uniforme. Considerando que esta CCI beneficia o GF, ela também pode ser classificada como DIF adversa. Como a probabilidade de respostas corretas é baixa, o item é considerado difícil.

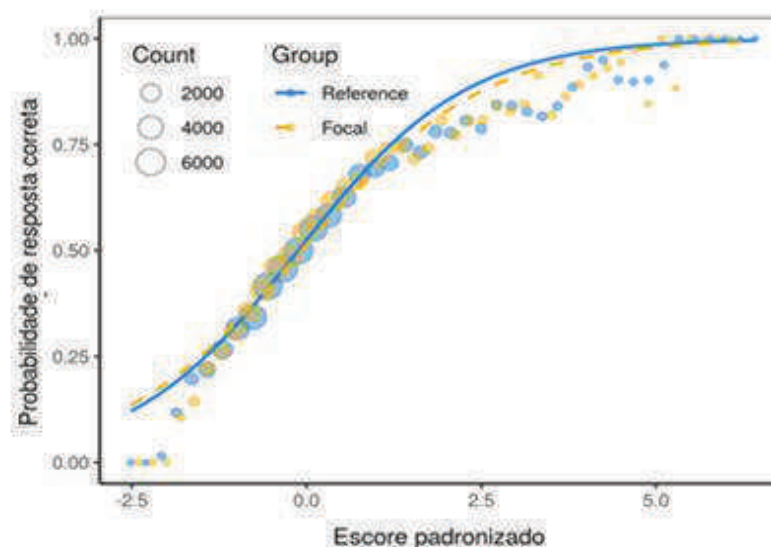


Gráfico 4-Questão 113 da 1ª aplicação do caderno azul 2018-2º dia de prova

No gráfico 4, é verificado um padrão análogo dos parâmetros dos itens tanto para o grupo de referência quando para o grupo focal, sugerindo DIF não significativo. As duas curvas se cruzam, deste modo o DIF é tipo não uniforme. Isto ocorre quando as CCI não são paralelas. O item, por ter probabilidade de acertos alta, é considerado como fácil.

Considerações finais

A partir desta pesquisa foi possível identificar que o nível de familiaridade, de acordo com a representação contida no enunciado, pode se apresentar como sendo mais próximo de um público específico (masculino ou feminino). Entretanto, embora um enunciado tenha elementos classificados como de maior ou menor aproximação para cada um deles, isso não pareceu influenciar no Funcionamento Diferencial do Item.

No que se refere o DIF do item 104, mesmo sendo categorizado como de média proximidade em relação ao contexto para ambos os grupos (focal e de referência), a probabilidade de respostas corretas em relação a habilidade ainda sim foi superior para os homens na situação investigada. Corroborando à suposição, o item 113 considerado mais próximo para as mulheres, e distante dos homens, não apresentou DIF significativa. Desde modo, a partir desse estudo, não é possível afirmar que elementos ou ações identificados como pertencentes em um universo gera impactos no DIF dos objetos analisados. Portanto, salienta-se a importância de continuar a investigação em outros itens pertencentes ao exame.

Na referida pesquisa, em razão das pessoas que participaram do preenchimento dos formulários não serem delimitadas por gênero e idade, para analisar o impacto de tais aspectos nos resultados, seria necessária uma segunda aplicação. Porém, com grupos determinados pela faixa etária e gênero. Deste modo, seriam coletadas informações para a comparação entre os dois grupos. A possibilidade será considerada em trabalhos futuros.

Referências

- ANDRIOLA, W. B. **Funcionamento Diferencial dos Itens (DIF)**: Estudo com Analogias para Medir o Raciocínio Verbal. *Psicologia: Reflexão e Crítica*, p. 475-483, 2000.
- BATESON, D. J.; PARSONS-CHATMAN, S. **Sex-related differences in science achievement**: A possible testing artifact. *International Journal of Science Education*, p. 371-385, 1989.
- BLASIS, E.; FALSARELLA, A. M.; ALAVARSE, O. M. **Avaliação e Aprendizagem**: Avaliações externas: perspectivas para a ação pedagógica e a gestão do ensino. Coordenação Eloisa de Blasis, Patricia Mota Guedes. – São Paulo: CENPEC: Fundação Itaú Social, 48 p., 2013.
- BRASIL. **Relatório pedagógico**: Enem 2011-2012. Brasília - DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2015. Disponível em: <<http://portal.inep.gov.br/documents/186968/484421/Relat%C3%B3rio+Pedag%C3%B3gico+Enem+2011-2012/b29257e3-2a6c-44a3-992a-02130c379ba9?version=1.2>>. Acesso em: 10/01/2020.

COLE, N. S. **History and development of DIF**. In: HOLLAND, P. W.; WAINER, H. (Ed.). *Differential item functioning*. Hillsdale: Lawrence Erlbaum, 1993.

DOUGLAS, J. A.; ROUSSOS, L. A.; STOUT, W. F. **Item-Bundle DIF hypothesis testing**: identifying suspect bundles and assessing their differential functioning. *Journal of Educational Measurement*, p. 465-484, 1996.

FACHINETTO, S. M. K. **Estudo comparativo entre o público EJA e o público regular utilizando a análise do Funcionamento Diferencial do item (DIF), na área de Ciências da Natureza e suas Tecnologias, a partir dos dados do Enem 2012**. Dissertação (mestrado profissional)-Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Métodos e Gestão em Avaliação, Florianópolis, 2016.

FERREIRA, E. A. **Teoria de reposta ao item – TRI** : análise de algumas questões do ENEM – habilidades 24 a 30, Mato Grosso do Sul: Dourados, 87 p., 2018.

MARCOM, G. S.; KLEINKE, M. U. **Gênero e Status socioeconômico**: reflexões sobre o desempenho dos candidatos na prova de Ciências da Natureza do Enem 2014. *Revista Perspectiva Sociológica*, 2017. Disponível em: <<https://cp2.g12.br/ojs/index.php/PS/article/view/1174/866>>. Acesso em: 15/02/2020.

MELLENBERGH, G. J. **Contingency table models for assessing item bias**. *Journal of Educational and Behavioral Statistics*, v. 7, n. 2, p. 105-118, 1982.

SUDWEEKS, R. R.; TOLMAN R. R. **Empirical Versus Subjective Procedures for Identifying Gender Differences in Science Test Items**. *Journal of research in science teaching*. Vol. 30, n. 1, p. 3-19, 1993.

VALLE, R. C. **Comportamento diferencial do item**: uma apresentação. *Estudos em Avaliação Educacional*, v. 25, p. 3-21, 2002.