

Análise de distratores: uma interpretação pelo modelo de resposta nominal dos itens do Enade 2017 aplicado à Licenciatura em Matemática

Analysis of distractors: an interpretation by the nominal response model of Enade 2017 items applied to the Degree in Mathematics

Análisis de distractores: una interpretación por el modelo de respuesta nominal de los ítems de Enade 2017 aplicados a la Licenciatura en Matemáticas

Sandra Cristina Martini Rostirola - Universidade do Estado de Santa Catarina | Instituto Federal Catarinense | Videira | Santa Catarina | Brasil. E-mail: sandra.rostirola@ifc.edu.br | Orcid: <https://orcid.org/0000-0003-2482-6117>

Elisa Henning - Universidade do Estado de Santa Catarina | Joinville | Santa Catarina | Brasil. E-mail: elisa.henning@udesc.br | Orcid: <https://orcid.org/0000-0002-7754-9451>

Ivanete Zuchi Siple - Universidade do Estado de Santa Catarina | Joinville | Santa Catarina | Brasil. E-mail: ivanete.siple@udesc.br | Orcid: <https://orcid.org/0000-0002-8640-1336>

Resumo: Testes baseados em itens de múltipla escolha frequentemente utilizam a Teoria da Resposta ao Item para a estimação de parâmetros dentro de modelos dicotômicos. Contudo, o Modelo de Resposta Nominal permite estudar itens politômicos, considerando toda a informação obtida na análise das alternativas escolhidas pelos respondentes. Nesse sentido, a presente pesquisa objetiva interpretar o conteúdo e as alternativas dos itens do Enade 2017 referentes aos cursos de Licenciatura em Matemática, a fim de contribuir para o aprofundamento da análise edumétrica do referido teste. Trata-se de uma abordagem quantiquantitativa que utiliza, concomitantemente, o Modelo de Resposta Nominal e a perspectiva da análise de erros em matemática. Os resultados indicam que o teste Enade 2017 para o curso de Licenciatura em Matemática continha itens com fragilidades em aspectos técnicos em termos de clareza, pertinência ou dificuldade, com necessidade de revisão por experts da área. Além disso, a análise desses itens permitiu explicar alguns raciocínios dos respondentes.

Palavras-chave: avaliação em larga escala; modelo de resposta nominal; licenciatura em matemática.

DOI: <http://dx.doi.org/10.1590/S1414-40772023000100042>

Abstract: Multiple-choice item-based tests often use Item Response Theory for parameter estimation within dichotomous models. However, the Nominal Response Model allows the study of polytomous items, considering all the information obtained from the analysis of alternatives chosen by respondents. In this regard, the present research aims to interpret the content and alternatives of the Enade 2017 items related to Mathematics Teaching courses to contribute to a deeper edumetric analysis of the said test. It is a quanti-qualitative approach that concurrently employs the Nominal Response Model and the perspective of mathematics error analysis. The results indicate that the Enade 2017 test for the Mathematics Teaching course contained items with weaknesses in technical aspects in terms of clarity, relevance, or difficulty, requiring review by experts in the field. Furthermore, the analysis of these items helped clarify some of the respondents' reasoning.

Keywords: large scale assessment; nominal response model; degree in mathematics.

Resumen: Las pruebas basadas en ítems de opción múltiple a menudo utilizan la Teoría de la Respuesta al Ítem para la estimación de parámetros dentro de modelos dicotómicos. Sin embargo, el Modelo de Respuesta Nominal permite estudiar ítems politómicos, considerando toda la información obtenida en el análisis de las alternativas elegidas por los encuestados. En este sentido, la presente investigación tiene como objetivo interpretar el contenido y las alternativas de los ítems del Enade 2017 relacionados con los cursos de Licenciatura en Matemáticas, con el fin de contribuir a un análisis edumétrico más profundo de dicho examen. Se trata de un enfoque cuanti-cualitativo que utiliza simultáneamente el Modelo de Respuesta Nominal y la perspectiva del análisis de errores en matemáticas. Los resultados indican que el examen Enade 2017 para el curso de Licenciatura en Matemáticas contenía ítems con debilidades en aspectos técnicos en términos de claridad, relevancia o dificultad, que requerían revisión por parte de expertos en el campo. Además, el análisis de estos ítems permitió explicar algunos razonamientos de los encuestados.

Palavras chave: avaliação a gran escala; modelo de resposta nominal; licenciatura en matemáticas.

1 Introdução

A Teoria da Resposta ao Item (TRI) fornece uma gama de ferramentas para medições de um determinado traço latente, as quais têm se mostrado precisas nas aplicações do Sistema de Avaliação da Educação Básica (SAEB) e do Exame Nacional do Ensino Médio (ENEM). No entanto, o Exame Nacional do Desempenho de Estudantes (ENADE) ainda não utiliza esses métodos, optando pela Teoria Clássica dos Testes (TCT). Na abordagem da TRI, tem-se o Modelo de Resposta Nominal (MRN) de Bock (1972), que considera todas as categorias de resposta de um item, utilizando toda a informação contida nas respostas dos indivíduos (Silva; Andrade; Borgatto; Nakamura, 2019).

Pinheiro, Costa e Cruz (2010) explicam que, no MRN, a alternativa mais indicada pelos sujeitos com maior traço latente é potencialmente a correta, enquanto as demais alternativas consistem em distratores. Se essa suposição não se cumpre, podemos considerar um erro de gabarito, problemas de digitação das respostas ou má formulação da questão, uma vez que o traço latente avaliado deve avançar de acordo com as respostas. Considerando esse raciocínio com a utilização do MRN, há possibilidades de analisar edumetricamente¹ uma determinada prova, salientando seus potenciais e fragilidades e conhecendo o raciocínio realizado pelos respondentes.

Desse modo, o uso desse modelo pode facilitar a interpretação da prova como um todo, uma vez que cada uma das alternativas é considerada na análise de um item. Isso pode representar um ganho qualitativo para uma prova como o Enade, por meio da qual se avaliam competências de concluintes do Ensino Superior brasileiro.

Este trabalho tem o objetivo de interpretar o conteúdo e as alternativas dos itens do Enade 2017 referentes aos cursos de Licenciatura em Matemática, a fim de

¹ Edumetria (ou variantes: edumétrico, edumetricamente) é o termo utilizado para expressar o aprofundamento da análise da qualidade de uma determinada avaliação, considerando sua validade e confiabilidade (Pereira; Oliveira; Tinoca, 2010).

contribuir para o aprofundamento da análise edumétrica do referido teste. Na metodologia quantiquantitativa adotada, apresenta-se um estudo que, além de estimar os parâmetros quantitativos a respeito dos itens, observa a estrutura destes sob uma perspectiva qualitativa, com o detalhamento de possíveis raciocínios dos respondentes dessa avaliação. Nesse sentido, defendemos a necessidade de se entender que o erro carrega informações que, muitas vezes, são desprezadas, considerando-se apenas a dicotomia certo e errado para formação de um escore. Com esse intuito, após a seção introdutória, informam-se os fundamentos teóricos utilizados, prosseguindo com o esclarecimento do método de pesquisa, a análise de dados e dos resultados obtidos e as considerações finais.

2 Revisão bibliográfica

As discussões sobre a avaliação educacional são de suma importância, uma vez que, por meio dela, é possível aferir, fundamentando-se em dados concretos, a qualidade da educação ofertada à população. Conforme Lopes e Vendramini (2015), a primazia dessas discussões se dá pela necessidade de um ensino substancialmente efetivo, haja vista seus efeitos nos mais diversos setores e camadas que compõem a organização sociopolítica e econômica de um país.

A TRI tem se tornado umas das principais ferramentas para análises de testes, especialmente aqueles que pertencem à esfera da avaliação educacional nos últimos anos, uma vez que permite uma medição precisa de um traço latente em termos uni e multidimensionais. Contudo, as pesquisas e aplicações mais comumente encontradas sugerem o uso de padrões de acerto ou erro, o que permite uma interpretação dicotômica dos itens do teste (Pinheiro; Costa; Cruz, 2010; Reise *et al.*, 2023; Smith; Bendjilali, 2022). Essa interpretação pressupõe a aplicação de um modelo para itens dicotômicos. Assim como nos modelos de Regressão Logística (Hosmer; Lemeshow, 2000), esse modelo binário pode ser adaptado para casos com mais de duas categorias,

ressalvadas as diferenças entre a análise de regressão e a TRI. Assim, é possível analisar itens politômicos, com escala ordinal ou nominal. A escala de medida é nominal se as categorias são puramente qualitativas e não houver ordenação natural (Agresti; Kateri, 2014; Kutner; Nachtsheim; Netter; Li, 2004). Nesse contexto, o Modelo de Resposta Nominal de Bock (1972) representa uma possibilidade para medições de itens de múltipla escolha considerados de forma politômica.

Bock (1972) assevera que, pelo MRN, é possível medir a probabilidade com que um indivíduo seleciona uma alternativa de resposta k de um determinado item i , como uma função dos parâmetros de dificuldades (c , na notação do MRN) e discriminação (a) associados ao traço latente do sujeito, conforme Equação (1):

$$P(u_i = x|\boldsymbol{\theta}; \mathbf{a}; \mathbf{c}) = \frac{e^{a_x(\theta) + c_x}}{\sum_{k=1,m} e^{a_{ik}(\theta) + c_{ik}}} \quad - \text{Equação (1)}$$

Tecnicamente, o procedimento para estimar os parâmetros por MRN não necessita da inserção do gabarito aos procedimentos de aplicação, como ocorre com os modelos logísticos de um, dois ou três parâmetros. Assim, os distratores, que se conceituam como alternativas diversas ao gabarito (Haladyna; Downing; Rodriguez, 2002), cumprem o papel de fornecer informação, pois a distância da resposta do respondente à alternativa correta traz suposições em relação à habilidade que se quer medir (Pinheiro; Costa; Cruz, 2010). Desse modo, o erro do respondente não o relega apenas a um escore reduzido, mas informa dados que dão indícios sobre a elaboração da questão e pode conduzir ao conhecimento do nível de desenvolvimento do respondente. Stewart *et al.* (2021) demonstram que um traço de habilidade está consistentemente associado ao pensamento correto, enquanto outras dimensões de habilidade representam modos específicos de pensamento incorreto.

Um importante estudo que converge para esta pesquisa é o de Thissen, Steinberg e Fitzpatrick (1989), no qual, em um primeiro momento, os autores explicam a evolução do Modelo de Bock (1972) na questão da superação de algumas dificuldades técnicas na estimação de parâmetros. Realizam, posteriormente, a análise de questões de múltipla escolha de uma avaliação de larga escala considerando o MRN e detalhando as alternativas distratoras, evidenciando possíveis raciocínios, níveis de habilidades e acerto ao acaso relacionados a cada item.

Além disso, estudos como os de Smith e Bendjilali (2022) e Zhang *et al.* (2021) explicam que a medição do traço latente no MRN incorpora quais respostas incorretas os respondentes selecionaram, reconhecendo, assim, que diferentes respostas incorretas podem indicar diferentes níveis de compreensão. Esses estudos apresentam uma alternativa para estimar parâmetros e computar escores nas avaliações em larga escala que pode ser uma ferramenta útil à análise dos próprios itens no que diz respeito a erros de formulação ou até mesmo às chamadas ‘pegadinhas’, itens que, em sua concepção, cumprem a função de distrair os respondentes de um teste. Assim, consideram-se, na análise, as conclusões de Beltrão e Mandarino (2023), os quais abordam que o formato do item influencia na dificuldade para a sua resolução e que alguns itens exigem um nível cognitivo mais elaborado do que outros menos exigentes quanto ao nível cognitivo.

Para a análise dos erros dos respondentes e para a interpretação dos possíveis raciocínios empregados nas resoluções, nos fundamentamos nos estudos de Allevato (2004), Cazorla (2002), Cury (2013), Cury e Cassol (2004) e Viali e Cury (2009). Também incorporamos no estudo a visão de Krutetskii (1976), que critica a psicometria enquanto apenas aferição de medidas para o certo e o errado, sem considerar o processo cognitivo subjacente. Dessa forma, defende-se a necessidade de se entender que o erro carrega informações que, muitas vezes, são desprezadas se consideramos

apenas a dicotomia certo e errado para formação de um escore. A próxima seção explica o método utilizado para análise dos itens do Enade 2017.

3 Metodologia

O estudo de abordagem quantiquantitativa analisa a prova do Enade 2017 para a Licenciatura em Matemática pelo Modelo de Resposta Nominal da TRI, o qual considera, para estimação dos parâmetros, as opções assumidas por um respondente em um teste de múltipla escolha.

Desse modo, a pesquisa busca realizar a Edumetria do Enade 2017, a qual permite uma análise efetiva dos itens e, ainda, um aprofundamento no entendimento do desenvolvimento do raciocínio dos respondentes (Pereira; Oliveira; Tinoca, 2010). Nesse sentido, tanto itens que estão respondidos corretamente como os que o indivíduo errou trazem informações pertinentes.

Os dados foram coletados nos Microdados do Enade 2017, disponibilizados pelo Ministério da Educação e Cultura (MEC)². Para o recorte amostral, foram considerados 10.869 participantes provenientes de cursos de Licenciatura em Matemática que tenham respondido, pelo menos, uma questão da prova de conhecimentos específicos.

A prova do Enade 2017 é composta por 40 itens, sendo 5 dissertativos e 35 objetivos. Na prova objetiva havia 8 itens de formação geral e 27 de formação específica, os quais estavam organizados em cinco alternativas (A, B, C, D, E), sendo apenas uma correta. Cada item da prova objetiva foi nomeado na análise pela letra "I" maiúscula, acompanhada de seu respectivo número na prova. O padrão de respostas (alternativas escolhidas) dos participantes para os 35 itens objetivos foi organizado em uma planilha analisada por meio do MRN com a utilização da interface do programa R – Rstudio (R Core Team, 2022) e dos pacotes Mirt (Chalmers, 2016) e Psych (Revelle,

² No sítio: <http://www.inep.gov.br>

2023). Os parâmetros para cada categoria de resposta foram estimados, sendo os dados complementados pelo uso da TCT.

Também foi calculada a proporção de escolha dos respondentes para cada alternativa em cada item, considerando, assim, o gabarito (resposta correta) e os distratores (outras alternativas presentes na questão). Conforme estudos de Martins (2018), os distratores podem ser classificados de acordo com a proporção de escolha. Um distrator forte possui uma porcentagem de escolha de 31% a 80%, um moderado de 21 a 30% e menores ou iguais a 20% são distratores fracos.

Dessa maneira, foram caracterizados os itens discrepantes pela observação das curvas gráficas e seus parâmetros quantitativos e realizada a análise edumétrica, que inclui a interpretação do raciocínio do respondente à luz da literatura específica da área da análise de erros em matemática. Isso permitiu o aprofundamento do conteúdo dos itens e a proposição de melhorias à prova, considerando sua importância na avaliação das competências do concluinte do curso de Licenciatura em Matemática no Brasil, conforme é exposto na próxima seção.

4 Resultados e discussões

4.1 Proporção de respondentes e estimação de parâmetros

Na Tabela 01 estão expressos os dados das proporções de respondentes para cada alternativa e para as respostas deixadas em branco (NA). São apresentadas em azul as proporções relativas à resposta correta e em vermelho destaca-se a ocorrência de distratores que possuem uma proporção maior de escolha do que o gabarito.

Tabela 01 – Proporção de distribuição das respostas por alternativa e não resposta (NA)

(em azul resposta correta, em vermelho distrator com probabilidade maior que o gabarito)

I (*)	Alternativas						I (*)	Alternativas					
	NA (**)	A	B	C	D	E		NA (**)	A	B	C	D	E
1	0,4	9,2	30,2	22,9	28,1	9,1	19	0,7	18,9	23,7	15,0	27,4	14,2
2	0,3	11,2	6,8	53,3	14,3	14	20	0,6	10,3	34,0	23,3	20,0	11,7
3	1,1	19,5	35,8	18,3	16,7	8,6	21	0,7	12,7	19,4	22,2	29,2	15,7
4	0,4	6,4	56,7	4,0	13,4	19,0	22	0,9	12,5	23,8	17,5	15,0	30,2
5	0,4	2,3	11,6	53,5	9,7	22,5	23	0,8	39,7	13,8	14,4	9,1	22,3
6	0,5	7,1	5,1	6,5	4,4	76,5	24	1,0	20,1	22,9	18,1	12,7	25,3
7	0,5	33	27,0	12,2	15,2	12,1	25	0,8	32,3	23,7	19,3	14,2	9,6
8	0,6	4,5	12,4	13,4	43,3	25,9	26	0,7	11,6	7,6	60,00	9,5	10,7
9	0,5	20,4	18,5	25,3	10,5	24,9	27	0,6	16,3	12,0	11,6	14,1	45,4
10	0,5	54,9	15,4	11,5	11,4	6,3	28	0,7	6,4	7,6	41,2	13,2	31,0
11	0,7	16,4	13,9	30,7	21,6	16,7	29	0,5	6,2	26,7	34,6	11,7	20,4
12	0,5	14,8	32,4	23,0	19,9	9,5	30	0,7	14,0	43,9	9,6	16,0	15,8
13	0,6	13,2	16,7	20,8	38,4	10,4	31	0,8	5,5	15,2	15,9	25,4	37,3
14	0,7	12,8	17,6	26,4	26,8	15,7	32	0,7	41,1	21,6	7,3	15,2	14,1
15	0,5	24,9	20,7	12,9	25,7	15,4	33	1,1	43,4	9,5	8,9	29,6	7,4
16	0,4	27,4	13,9	32,3	11,1	14,8	34	0,9	14,4	19,7	17,7	34,7	12,6
17	0,7	25,2	18,7	27,0	18	10,5	35	0,8	9,5	7,2	9,2	65,3	7,9
18	0,9	17,5	21,9	22,5	20,9	16,2	-	-	-	-	-	-	-

*: Nomeação do item; **: Não alternativa.

Fonte: Elaboração própria (2023).

Os itens, em sua generalidade, apresentam a maior porcentagem de respondentes na alternativa que representa o gabarito, conforme Tabela 01. Porém, os itens: I1, I14, I16, I18, I20, I23 e I24 apresentam maior porcentagem de respondentes em um distrator. Os distratores apresentados como maior proporção de escolha na Tabela 01 são classificados como moderados (I14, I18 e I24) a fortes (I1, I16, I20 e I23).

Os parâmetros relativos a cada uma das categorias de respostas possíveis (A, B, C, D, E) foram estimados pelo MRN, resultando em probabilidades descritas como a_1 , a_2 , a_3 , a_4 , a_5 na Tabela 02. Os números positivos indicam respostas mais prováveis e os negativos indicam valores associados a respostas menos prováveis (Thissen; Steinberg; Fitzpatrick, 1989).

Tabela 02 – Estimativas dos parâmetros por alternativa das questões obtidas pelo MRN

I	a_1	a_2	a_3	a_4	a_5	I	a_1	a_2	a_3	a_4	a_5
I1	0,208	-0,324	0,523	-0,223	-0,183	I19	-0,133	0,171	-0,2	0,129	0,033
I2	-0,09	-0,259	0,201	0,069	0,079	I20	-0,088	-0,044	0,27	-0,2 0	0,062
I3	-0,012	0,300	-0,269	-0,163	0,144	I21	-0,055	0,094	-0,159	-0,005	0,126
I4	-0,28	0,844	-0,546	-0,047	0,028	I22	-0,098	-0,138	-0,222	-0,147	0,6
I5	-0,454	0,01	0,594	-0,087	-0,063	I23	0,202	-0,182	-0,067	-0,35	0,396
I6	-0,178	-0,229	-0,422	-0,105	0,93	I24	-0,042	0,316	-0,177	-0,207	0,111
I7	0,305	-0,04	-0,289	0,273	-0,25	I25	0,41	0,235	0,199	-0,244	-0,203
I8	-0,458	-0,131	-0,049	0,51	0,127	I26	-0,18	-0,418	0,793	-0,269	0,073
I9	-0,013	-0,2	0,261	-0,407	0,359	I27	0,008	-0,316	-0,282	-0,068	0,658
I10	0,654	-0,015	-0,327	-0,263	0,04	I28	-0,466	-0,632	0,66	-0,078	0,516
I11	-0,305	-0,36	0,422	0,008	0,235	I29	-0,662	0,075	0,46	-0,118	0,244
I12	-0,141	0,572	-0,269	-0,257	0,095	I30	-0,212	0,72	-0,596	0,001	0,087
I13	-0,524	-0,583	0,23	0,595	0,282	I31	-0,466	-0,306	0,1	0,146	0,52
I14	0,038	0,158	-0,113	0,014	-0,097	I32	0,607	0,101	-0,527	-0,333	0,152
I15	-0,08	-0,331	-0,115	0,393	0,133	I33	0,286	-0,563	-0,434	0,207	0,504
I16	0,125	-0,023	-0,11	0,052	-0,045	I34	-0,139	-0,066	-0,21	0,358	0,057
I17	-0,014	0,107	0,287	-0,068	-0,313	I35	-0,25	-0,454	0,031	0,698	-0,025
I18	-0,017	-0,157	-0,07	-0,121	0,365	-	-	-	-	-	-

Fonte: Elaboração própria (2023).

Na análise das cinco categorias, cujas probabilidades de escolha são representadas pelos parâmetros a_1 , a_2 , a_3 , a_4 e a_5 na Tabela 02, é esperado que a categoria a_3 relativa ao gabarito, tenha o maior valor entre os parâmetros. Além disso, entende-se, pelo MRN, que a curva relacionada ao gabarito é ascendente em relação à proficiência, ou seja, tem sua probabilidade de acerto aumentada quando θ – o traço latente do respondente, que pode ser entendido como uma medida de aptidão ou proficiência - for maior (Pasquali, 2018).

No entanto, ocorreram itens nos quais a resposta correta (gabarito) apresentava resultados discrepantes em relação aos seus distratores, o que poderia ser observado também nas curvas de alguns itens. O cálculo de θ no MRN incorpora quais respostas incorretas os alunos selecionam, reconhecendo, assim, que diferentes respostas incorretas podem indicar diferentes níveis de compreensão (Smith; Bendjilali, 2022).

³ a_s : é a denominação das categorias no Modelo de Resposta Nominal.

A próxima seção se reporta às análises quantitativas do item seguidas de sua interpretação qualitativa.

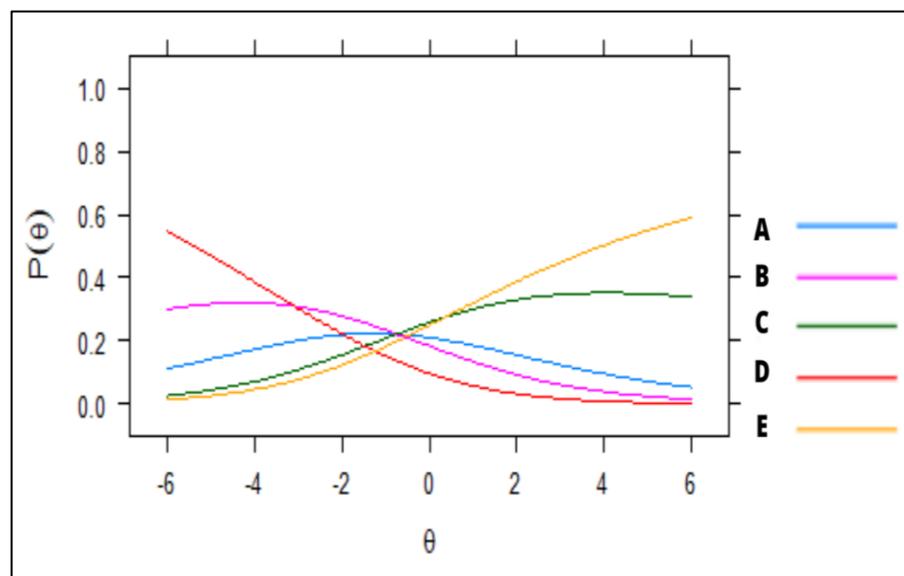
4.2 Perspectiva quanti-qualitativa para avaliação dos itens

A estimação dos parâmetros dos itens pelo MRN resultou em achados similares aos estudos de Thissen, Steinberg e Fitzpatrick (1989), ou seja, itens que apresentam parâmetros maiores e positivos associados à categoria correta ($a_{correct}$), enquanto encontramos escores menores ou negativos nas categorias distratoras. Nesse sentido, vamos particularizar os itens I9, I16, I19, I21 e I23, que se caracterizam por distratores com parâmetros e curvas que escapam ao padrão do modelo, e analisá-los na perspectiva da análise de erros em matemática.

O item 09 possui gabarito C como parâmetro $a_3=0,261$, porém foi estimada probabilidade maior para o distrator E ($a_5= 0,359$). Na Figura 01⁴, percebemos que a curva A estabiliza à medida que o θ aumenta, ficando abaixo de E. Salientamos que a Tabela 01 mostra proporções aproximadas equivalentes entre o gabarito e o distrator E.

⁴ Os gráficos que representam as curvas dos itens, encontrados nas Figuras 01, 03, 05, 07 e 09 podem apresentar curvas crescentes e decrescentes, relacionadas a maior ou menor probabilidade de escolha de uma alternativa pelo respondente. A alternativa correta é usualmente escolhida por indivíduos com proficiência maior. A escala de aptidão/proficiência no gráfico está organizada de -6 a +6, que é o padrão da apresentação no R, proporcionando melhor visualização da amplitude dos escores.

Figura 01 – Gráfico da Curva do item 09



Fonte: Elaboração própria (2023).

Esses dados permitem uma análise qualitativa do item (Figura 02) que trata do conteúdo de integrais. Nessa questão, era preciso reconhecer a aplicação da integral definida como a área de regiões delimitadas por funções, mas também identificar que o valor numérico da integral de função negativa é negativo e, portanto, no item II, o resultado da integral não representa o valor da área da região delimitada pelas curvas. O valor igual a 3 indicado na afirmação II atuou como um distrator, atraindo os respondentes para a alternativa errada. Para Cury (2013), sem o domínio da técnica, como habilidades de lidar com regras para cálculo de limites, derivadas ou integrais, por exemplo, o aluno não tem ferramentas para trabalhar com os conceitos. Isso parece estar prejudicando a correta aferição de resposta ao item 09.

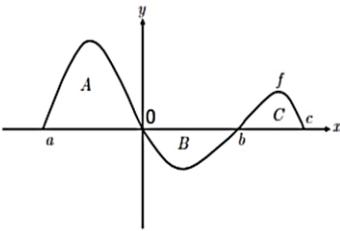
Figura 02 – Item 09 do Enade 2017 da prova de Licenciatura em Matemática

Considere $f : [a, c] \rightarrow \mathbb{R}$ uma função contínua e $b \in (a, c)$, conforme ilustra o gráfico abaixo. Represente por:

A a área da região limitada pela reta de equação $y = 0$ e pela curva $\{(x, f(x)); x \in [a, 0]\}$;

B a área da região limitada pela reta de equação $y = 0$ e pela curva $\{(x, f(x)); x \in [0, b]\}$;

C a área da região limitada pela reta de equação $y = 0$ e pela curva $\{(x, f(x)); x \in [b, c]\}$.



Sabendo-se que $A = 5$, $B = 3$ e $C = 2$, avalie as afirmações a seguir.

I. $\int_a^0 f(x) dx = 5$

II. $\int_0^b f(x) dx = 3$

III. $\int_a^c f(x) dx = 4$

É correto o que se afirma em

A I, apenas.

B II, apenas.

C I e III, apenas.

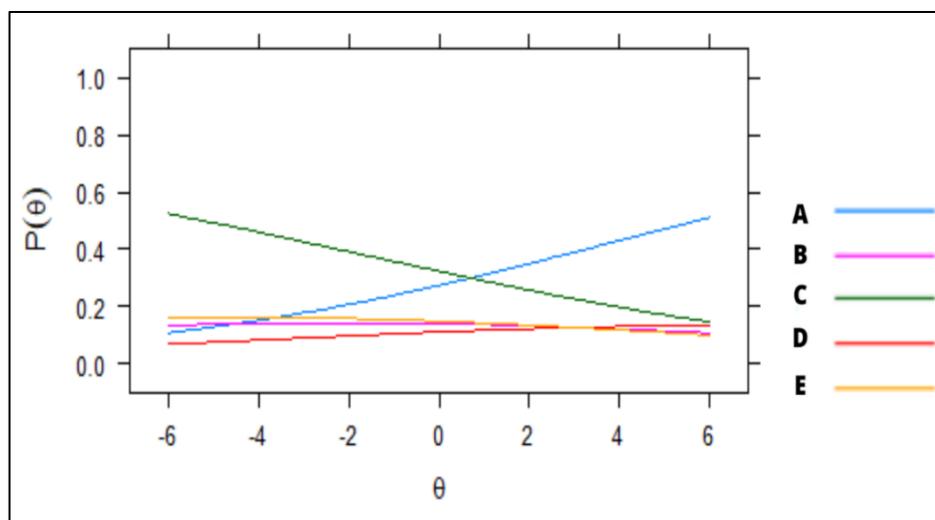
D II e III, apenas.

E I, II e III.

Fonte: Adaptada da Prova Enade (2017).

As estimativas para o item 16 indicam a probabilidade para o gabarito – letra D, em 0,052 (a_4), enquanto o distrator A apresenta probabilidade de 0,125 (a_1). A Tabela 01 ainda traz a informação de que a proporção de escolha do gabarito foi inferior à de A (11,1% e 27,4%, respectivamente). Contudo, a maior escolha ainda se deu para o distrator C (32,6%), o que amplifica as distorções para esse item. Graficamente, a Figura 03 mostra que a curva A relativa à alternativa A cresce em relação ao aumento do θ , mantendo-se com probabilidade baixa para o gabarito (D).

Figura 03– Curva do item 16



Fonte: Elaboração própria (2023).

No item 16, o erro decorre de o respondente realizar a substituição dos valores nas variáveis no sistema, não analisando se o problema possui outras soluções e não relacionando aos fundamentos matemáticos de álgebra linear ou geometria analítica. Conforme Cury (2013), esse é um tipo de erro que indica que os respondentes de um teste não entendem o processo que deve ser realizado e partem das informações que já possuem para deduzir o que deveria ser feito.

Figura 04– Item 16 do Enade 2017 da prova de Licenciatura em Matemática

A solução de um sistema linear de três equações e três incógnitas pode ser interpretada geometricamente como a interseção de três planos no espaço e consiste em verificar se os três planos têm um único ponto, infinitos pontos ou nenhum ponto em comum, para determinar se o sistema possui solução única, infinitas soluções ou nenhuma solução, respectivamente.

Com base nessas informações, conclui-se que o sistema linear

$$\begin{cases} x+2y+z = 1 \\ y+z = 2 \\ x+y = -1 \end{cases}$$

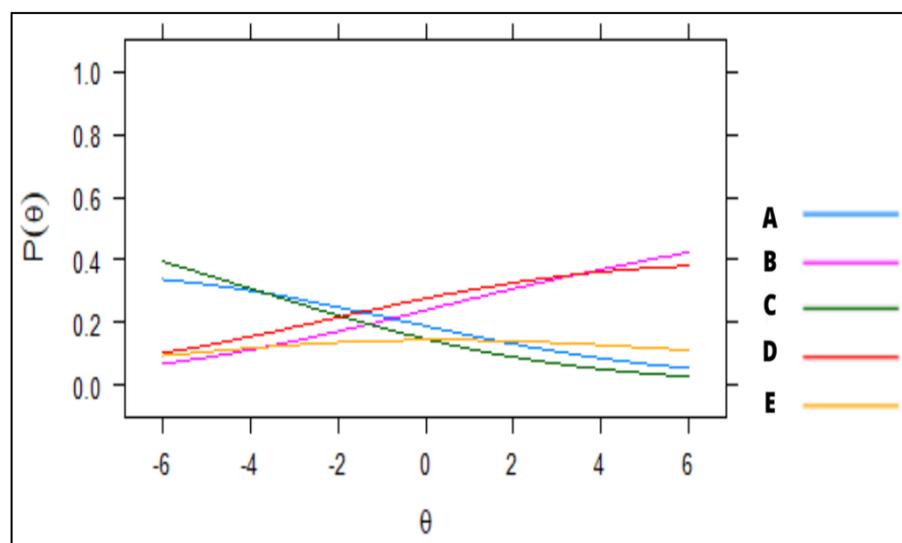
tem como solução

- A** o ponto $(0,-1,3)$.
- B** o plano que passa pelo ponto $(0,-1,3)$ e que possui como vetor normal o vetor $\vec{n} = (1,2,1)$.
- C** a reta que passa pelo ponto $(0,-1,3)$ e que possui como vetor diretor o vetor $\vec{v} = (1,2,-1)$.
- D** a reta que passa pelo ponto $(0,-1,3)$ e que possui como vetor diretor o vetor $\vec{v} = (1,-1,1)$.
- E** o conjunto vazio.

Fonte: Prova Enade (2017).

Com relação à questão 19, ocorreram dois distratores com probabilidade positiva (B com $a_2 = 0,171$ e E com $a_5 = 0,033$). O gabarito (D) obteve probabilidade de $0,129$ (a_4), sendo, dessa forma, inferior ao distrator B. Na Figura 05, observamos a curva relativa à alternativa D decrescendo suavemente em relação à B.

Figura 05– Curva do Item 19



Fonte: Elaboração própria (2023).

Na resolução da questão (Figura 06), podemos inferir que a grande quantidade de passos para resolução pode ter levado o respondente a uma interpretação equivocada dos primeiros resultados: o respondente precisava interpretar geometricamente a situação dada, sendo que, para isso, deveria transitar entre as representações numérica e gráfica dos pontos dados (que representavam as cidades). Importa, aqui, explorar Duval (2012), a quem o recurso a muitos registros parece uma condição necessária para que os objetos matemáticos não sejam confundidos e que possam também ser reconhecidos em cada uma de suas representações. A fragilidade nessa transição de representações e nos conhecimentos de geometria analítica pode ter contribuído para tais resultados apresentados pelos respondentes. Esse aspecto também é analisado por Cury (2013), que fala de classes de erros matemáticos em que o respondente não consegue finalizar corretamente o raciocínio desenvolvido pela incompreensão de algumas subetapas.

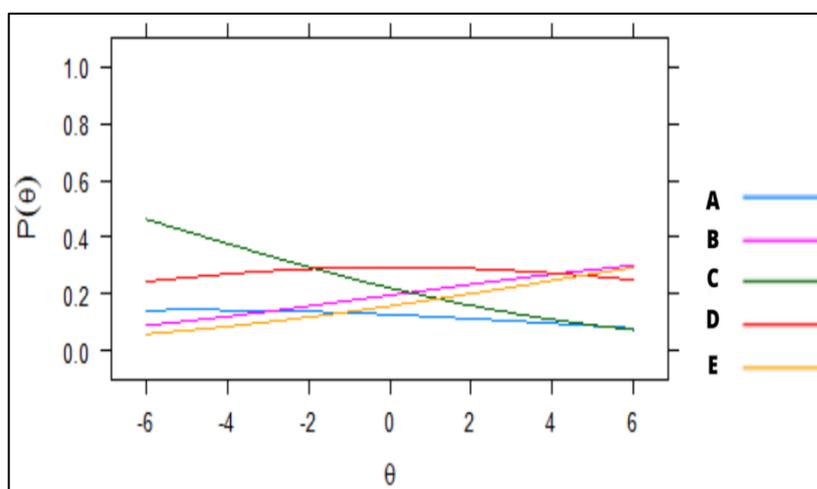
Figura 06 – Item 19 do Enade 2017 da prova de Licenciatura em Matemática

<p>De acordo com a Política Nacional de Resíduos Sólidos, a implantação da coleta seletiva é obrigação dos municípios, e metas referentes à coleta seletiva fazem parte do conteúdo mínimo que deve constar nos planos de gestão integrada de resíduos sólidos dos municípios.</p> <p>Disponível em: <http://www.mma.gov.br>. Acesso em: 10 jul. 2017 (adaptado).</p> <p>Suponha que os prefeitos de três cidades vizinhas decidiram fazer um projeto conjunto para a construção de um armazém para separação de materiais recicláveis e que pudesse funcionar também como um ponto de entrega voluntária. Os prefeitos decidiram que o armazém deveria ser construído em um local equidistante aos centros das cidades, representados em um plano cartesiano pelos pontos $A = (1, 3)$, $B = (10, 0)$ e $C = (3, 7)$.</p>	<p>Nessa situação, quais são as coordenadas do ponto escolhido?</p> <p>A $(5, 3)$</p> <p>B $\left(\frac{11}{2}, \frac{7}{2}\right)$</p> <p>C $\left(6, \frac{5}{2}\right)$</p> <p>D $(6, 3)$</p> <p>E $\left(\frac{13}{2}, \frac{7}{2}\right)$</p>
---	--

Fonte: Prova Enade (2017).

O item 21 possui os distratores B e E com probabilidade ($a_2 = 0,094$ e $a_5 = 0,126$), enquanto o gabarito possui probabilidade $-0,005$ (a_4). Quando analisado pela TCT, apresentou bisseriais negativos para todas as alternativas, sendo o maior valor referente ao gabarito ($-0,01$), mas muito próximos dos distratores B e E ($-0,07$ e $-0,05$ respectivamente). A curva D (Figura 07) decrescente, conforme avança o θ , confirma as incoerências quantitativas apresentadas.

Figura 07 – Curva do Item 21



Fonte: Elaboração própria (2023).

Na resolução de I21 (Figura 08) é preciso analisar alguns aspectos quanto à soma das probabilidades da afirmação I. Os primeiros cálculos chegam a três frações iguais a $1/15$. Se não ocorrer a interpretação de que é preciso efetuar a adição, erroneamente se considera a alternativa E como verdadeira. Essa discrepância pode ser observada na Tabela 06, que indica a proporção de 30% dos participantes optando por esse distrator. Em consonância a Viali e Cury (2009), esse erro pode ser considerado uma subclasse de uma situação em que o respondente conhece o conceito de probabilidade como o quociente entre o número de casos favoráveis e o número de casos possíveis, mas evidencia uma dificuldade em discernir eventos compostos.

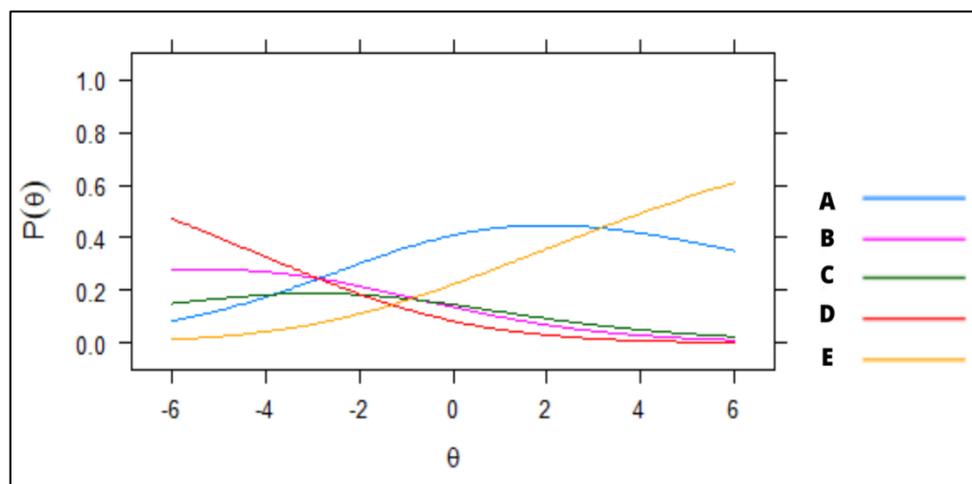
Figura 08– Item 21 do Enade 2017 da prova de Licenciatura em Matemática

<p>Seis estudantes se inscreveram para um campeonato escolar de xadrez: três meninas, das quais duas são irmãs gêmeas, e três meninos. Na primeira rodada serão formadas as três duplas de adversários por sorteio, da seguinte forma: o primeiro jogador é sorteado entre os seis participantes; o segundo é sorteado entre os cinco restantes; o terceiro entre os quatro restantes; o quarto, entre os três restantes; a primeira dupla é formada pelo primeiro e segundo sorteados; a segunda dupla é formada pelo terceiro e quarto sorteados; a terceira dupla é formada pelos dois últimos que não foram sorteados.</p> <p>Considerando essas condições a respeito da formação das duplas de adversários na primeira rodada do campeonato, avalie as afirmações a seguir</p>	<p>I. A probabilidade de as gêmeas se enfrentarem é de $\frac{1}{15}$.</p> <p>II. A probabilidade de a primeira dupla sorteada ser de meninos é de $\frac{1}{5}$.</p> <p>III. A probabilidade de a primeira dupla sorteada ser composta por uma menina e um menino é de $\frac{3}{5}$.</p> <p>É correto o que se afirma em</p> <p>A I, apenas.</p> <p>B II, apenas.</p> <p>C I e III, apenas.</p> <p>D II e III, apenas.</p> <p>E I, II e III.</p>
---	--

Fonte: Prova Enade (2017).

Para o item 23 (Figura 09) foram estimadas, pelo MRN, probabilidades 0,396 (a_5) para o gabarito E, enquanto o distrator A ficou com 0,202 (a_1). Também com relação à opção dos respondentes, 20% optaram pela letra A e 40% pela letra B. No entanto, apresentou, pela avaliação por TCT, bisseriais positivos para a alternativa A (0,01) e para o gabarito E (0,03). O modelo gráfico mostra que as curvas B, C e D reduzem a probabilidade conforme avança o valor de θ . Já a curva A tem um ligeiro aumento em alguns valores de θ , sendo superada pelo gabarito, mas indicando a escolha do distrator por respondentes com níveis altos de habilidade.

Figura 09– Curva do item 23



Fonte: Elaboração própria (2023).

Diante desses aspectos, o conteúdo dos distratores do item 23 (Figura 10) mostrou que os respondentes são atraídos pelo distrator A, uma vez que realizados os cálculos pertinentes encontramos o valor discriminado $-33/2$. No entanto, há um ponto de descontinuidade na função, indicando que 0 faz parte do intervalo $[-2,1]$, estabelecendo uma função indefinida, sendo validado o gabarito E. Cury e Cassol (2004) nos falam sobre erros relativos à aprendizagem de cálculo e de erros que envolvem o intervalo numérico que faz parte da situação-problema.

Figura 10 – Item 23 do Enade 2017 da prova de Licenciatura em Matemática

Considerando que um estudante esteja testando um *software* para calcular o valor da integral $\int_{-2}^1 \left(\frac{1}{x^2} - 5 \right) dx$, avalie as asserções a seguir e a relação proposta entre elas.

I. O resultado $\int_{-2}^1 \left(\frac{1}{x^2} - 5 \right) dx = -\frac{33}{2}$, apresentado pelo *software*, está correto.

PORQUE

II. A primitiva da função $f(x) = \frac{1}{x^2} - 5$ é a função $F(x) = -\frac{1}{x} - 5x$ e, pelo Teorema Fundamental do Cálculo, conclui-se que

$$\int_{-2}^1 \left(\frac{1}{x^2} - 5 \right) dx = \left(-\frac{1}{x} - 5x \right) \Big|_{-2}^1 = \left(-\frac{1}{1} - 5 \cdot 1 \right) - \left(-\frac{1}{(-2)} - 5(-2) \right) = -\frac{33}{2} .$$

A respeito dessas asserções, assinale a opção correta.

A As asserções I e II são proposições verdadeiras, e a II é uma justificativa correta da I.
B As asserções I e II são proposições verdadeiras, mas a II não é uma justificativa correta da I.
C A asserção I é uma proposição verdadeira, e a II é uma proposição falsa.
D A asserção I é uma proposição falsa, e a II é uma proposição verdadeira.
E As asserções I e II são proposições falsas.

Fonte: Adaptada da Prova Enade (2017).

Assim, dentro das perspectivas dos 35 itens analisados, temos 5 (14%) que apresentam fragilidades na construção de suas alternativas (distratores), ressaltando que os itens I16, I19 e I21 e I23 foram retirados do cômputo das notas pela sistemática adotada pelo Enade 2017 (Brasil, 2018), porque apresentaram índice de discriminação (correlação ponto bisserial) muito baixo. Contudo, o estudo os considerou, pois efetivamente foram respondidos pelos participantes do Enade 2017. Assim, é preciso entender tecnicamente também esses itens.

Com relação a essa retirada de itens pelo Enade 2017, a Matriz de Referência (Brasil, 2017), com os itens I16 e I19, propunha avaliar a habilidade de “Formular conjecturas e generalizações estabelecendo relações entre os aspectos formais e intuitivos”, permanecendo apenas um item que possui essa expectativa. Quanto ao item I21, a habilidade mensurada era “Resolver problemas”, contudo, a retirada desse item fragiliza a prova, haja vista que apenas dois itens (I20 e I21) pretendiam avaliar o objeto de conhecimento probabilidade e estatística. Além disso, o I23 é o único que traz elementos do objeto de conhecimento fundamentos de análise, criando uma

lacuna avaliativa importante nesse aspecto. Quanto a I9, na fase de elaboração da prova, seria admissível ser reavaliado por especialistas, uma vez que apresenta problemas em seus distratores.

Nesta análise, os dados refletem que a sistemática adotada pelo INEP para o Enade interfere na avaliação das competências do licenciando em matemática. Dessa maneira, Alves (2020) indica que não se deve partir da eliminação de itens, mas apontar a necessidade de consultar especialistas na intenção de avaliar se a presença desses itens é essencial à validade de conteúdo e se a recomendação de eliminação será o caminho mais parcimonioso e plausível.

5 Considerações finais

Este estudo tinha por objetivo interpretar o conteúdo e alternativas dos itens do Enade 2017 referentes aos cursos de Licenciatura em Matemática, a fim de contribuir para o aprofundamento da análise edumétrica do referido teste. Nesse sentido, não era de interesse específico trabalhar um cenário de parametrização da TRI, por meio do modelo logístico de 3 parâmetros, mas obter dados que permitissem avaliar não somente os aspectos técnicos de formulação dos itens, como também observar possíveis raciocínios que implicam na compreensão de lacunas na formação desses futuros professores, o que é determinante para ações didático-pedagógicas nos cursos dentro da esfera da Educação Matemática.

Nesse âmbito, a análise do teste partiu do geral para o particular, identificando possíveis discrepâncias no comportamento dos parâmetros de cada uma das alternativas. Esperava-se que os parâmetros alcançassem valores maiores à medida que o traço latente avançava, situação que poderia ser observada nos gráficos das curvas dos itens. Contudo, a proporção de 5/27 desses itens apresentaram-se diferentes da abordagem esperada, indicando que o teste Enade 2017 tinha itens cujos distratores deveriam ter sido melhor analisados em seu aspecto técnico. Além disso,

revelam fragilidades no raciocínio matemático dos estudantes, sugerindo dificuldades na compreensão semiótica associada à linguagem matemática e dificuldades em técnicas de cálculo e de álgebra.

Por meio da aplicação do Modelo de Resposta a Itens Nominal (MRN), torna-se viável a criação de representações gráficas que apontam opções mais atrativas para os estudantes. Nesse modelo, também é possível identificar as respostas mais frequentemente escolhidas pelos alunos, adaptadas aos seus níveis de proficiência. Isso possibilita uma análise abrangente que contempla tanto a totalidade quanto as nuances individuais. Dentro dessa abordagem, também é factível estimar os parâmetros de discriminação, os quais diferenciam os valores associados às respostas mais prováveis, conforme abordado por Reise *et al.* (2023). O MRN fornece um alicerce consistente para uma análise que combina elementos quantitativos e qualitativos, permitindo observações detalhadas sobre a evolução dos estudantes e suas capacidades. Além disso, esse modelo é capaz de explicar a escolha de alternativas menos corretas, revelando até mesmo palpites aleatórios. Essa capacidade é de suma importância para identificar pontos críticos e lacunas no conhecimento do conjunto de respondentes. Assim, possibilita uma análise suplementar de natureza qualitativa, ao rastrear possíveis equívocos dos estudantes e conectá-los às investigações relacionadas à análise de erros.

Os resultados apontam, ainda, que é preciso questionar os entes administrativos quanto à experimentação de novas sistemáticas que possam dar maior precisão aos resultados. Estes, por sua vez, implicam na recepção, pela sociedade, de um educador com saberes potenciais para suas vivências profissionais, o que amplia a oportunidade da oferta de educação de qualidade aos cidadãos.

Nesse âmbito, é preciso enfatizar as abordagens de Vianna (2009), o qual defende que os resultados de avaliações educacionais não devem ser usados única e exclusivamente para traduzir um certo desempenho. Eles podem implicar na definição

de novas políticas públicas, de projetos de implantação e modificação de currículos, de programas de formação continuada dos docentes e, de maneira decisiva, na definição de elementos para a tomada de decisões que visem a provocar um impacto, ou seja, mudanças no pensar e no agir dos integrantes do sistema. Desse modo, os resultados obtidos via MRN têm a potencialidade de melhorar não apenas a construção técnica da prova, mas também podem impactar em mudanças nas diretrizes da avaliação Enade e na estrutura curricular dos cursos de formação inicial.

Agradecimentos

Agradecemos à Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina – FAPESC (Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina, financiamento nº 2023TR000334); ao Programa de Bolsas Universitárias de Santa Catarina – UNIEDU e ao grupo de pesquisa NEPESTEEM.

Referências

AGRESTI, A.; KATERI, M. Categorical data analysis. *In*: LOVRIC, M. (ed.) **International encyclopedia of statistical science**. Heidelberg: Springer, 2011. p. 206–208. Disponível em: https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2_161. Acesso em 20 out. 2023.

ALLEVATO, N. S. G. Resolução de problemas, software gráfico e detecção de lacunas no conhecimento da linguagem algébrica. *In*: ENCONTRO NACIONAL DE EDUCAÇÃO MATEMÁTICA - ENEM, 8., 2004, Recife. **Anais [...]**. Recife: SBEM, 2004. p. 1-20. Disponível em: <http://www.sbembrasil.org.br/files/viii/pdf/06/CC47973757953.pdf>. Acesso em: 07 out. 2023.

ALVES, C, K. T. **Análise das propriedades psicométricas da prova de conhecimentos específicos de licenciatura em ciências biológicas no ENADE 2017**. 2020. Dissertação (Mestrado Profissional em Métodos e Gestão em Avaliação) - Universidade Federal de Santa Catarina, Florianópolis, 2020. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/216664/PMGA0048-D.pdf?sequence=-1&isAllowed=y>. Acesso em: 16 maio 2023.

BELTRÃO, K. I.; MANDARINO, M. C. F. Análise dos itens de múltipla escolha das provas do Enade 2016. **Estudos em Avaliação Educacional**, São Paulo, v. 34, p. e07951, 2023. Disponível em:

<https://publicacoes.fcc.org.br/eae/article/view/7951>. Acesso em: 29 ago. 2023.

BOCK, R. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, Heidelberg, v. 37, n. 1, p. 29-51, 1972. Disponível em: <https://link-springer-com.ez74.periodicos.capes.gov.br/article/10.1007/BF02291411>. Acesso em: 07 out. 2023.

BRASIL. Instituto Nacional de Pesquisas e Estudos Educacionais Anísio Teixeira. **Relatório síntese de área**: matemática. Brasília: Inep, 2017. Disponível em: https://download.inep.gov.br/educacao_superior/enade/relatorio_sintese/2017/Matematica.pdf. Acesso em: 16 maio 2023.

CAZORLA, I. M. **A relação entre a habilidade viso-pictórica e o domínio de conceitos estatísticos na leitura de gráficos**. 2002. Tese (Doutorado em Educação) - Universidade de Campinas, Campinas, 2002. Disponível em: https://www.psiem.fe.unicamp.br/pf-psiem/cazorla_irenemauroicio_d.pdf. Acesso em: 16 maio 2023.

CHALMERS, R. P. Generating adaptive and non-adaptive test interfaces for multidimensional item response theory applications. **Journal of Statistical Software**, EUA, v. 71, p. 1-38, 2016. Disponível em: <https://www.jstatsoft.org/article/view/v071i05>. Acesso em: 07 out. 2023.

CURY, H. N. **Análise de erros**: o que podemos aprender com as respostas dos alunos. Belo Horizonte: Autêntica, 2013.

CURY, H. N.; CASSOL, M. Análise de erros em cálculo: uma pesquisa para embasar mudanças. **Acta Scientiae**: revista de ensino de ciências e matemática, São Paulo, v. 6, n. 1, p. 27-36, 2004. Disponível em: <http://www.periodicos.ulbra.br/index.php/acta/article/view/128/116>. Acesso em: 07 out. 2023.

DUVAL, R. THADEU, M. (tradução). Registros de representação semiótica e funcionamento cognitivo do pensamento. **Revemat**: revista eletrônica de matemática, Santa Catarina, v. 7, n. 2, p. 266-297, 2012. Disponível em: <https://periodicos.ufsc.br/index.php/revemat/article/view/1981-1322.2012v7n2p266>. Acesso em: 07 out. 2023.

HALADYNA, T. M.; DOWNING, S. M.; RODRIGUEZ, M. C. A review of multiple-choice item-writing guidelines for classroom assessment. **Applied Measurement in Education**, São Paulo, v. 15, n. 3, p. 309-333, 2002. Disponível em: <http://site.ufvjm.edu.br/fammuc/files/2016/05/item-writing-guidelines.pdf>. Acesso em: 07 out. 2023.

HOSMER, D. W.; LEMESHOW, S. **Applied logistic regression**. Wiley: New York, 2000.

KRUTETSKII, V. A. **The psychology of mathematical abilities in schoolchildren**. Chicago: University of Chicago Press, 1976.

KUTNER, M. H.; NACHTSHEIM, C. J.; NETTER, J.; LI, W. **Applied linear statistical models**. Boston: McGraw-Hill Irwin, 2005. Disponível em: https://users.stat.ufl.edu/~winner/sta4211/ALSM_5Ed_Kutner.pdf. Acesso em: 07 out.

LOPES, F. L.; VENDRAMINI, C. M. M. Propriedades psicométricas das provas de pedagogia do Enade via TRI. **Avaliação: revista da avaliação da educação superior**, Campinas; Sorocaba, v. 20, n. 1, p. 27-47, mar. 2015. Disponível em: <https://www.scielo.br/j/aval/a/64h6cvwZrKPR9v4P4nChZ6t/?format=pdf&lang=pt>. Acesso em: 07 out. 2023.

MARTINS, V. B. **Habilidades de estudantes do ensino médio para planificar figuras tridimensionais**. 2018. Monografia (Especialização em Ciências e Tecnologias da Educação) - Instituto Federal Sul – Riograndense, Pelotas, 2018. Disponível em: <https://viniciuscavg.files.wordpress.com/2018/06/2018-vanderlei-planificac3a7c3a3o.pdf>. Acesso em: 16 maio 2023.

PASQUALI, L. **TRI – teoria da resposta ao item: teoria, procedimentos e aplicações**. Curitiba: Apris, 2018.

PEREIRA, A.; OLIVEIRA, I.; TINOCA, L. A cultura de avaliação: que dimensões. *In*: ENCONTRO INTERNACIONAL TIC E EDUCAÇÃO: TICeduca, 1., 2010, Lisboa. **Actas** [...]. Lisboa, 2010. p. 350-357.

PINHEIRO, I. R.; COSTA, F. R.; CRUZ, R. M. Modelo nominal da teoria de resposta ao item: uma alternativa. **Avaliação Psicológica**, São Paulo, v. 9, n. 3, p. 437-447, 2010. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000300010. Acesso em: 07 out. 2023.

R CORE TEAM. **R: a language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, Áustria, 2022. Disponível em: <https://www.R-project.org/>. Acesso em: 17 maio 2023.

REISE, S. P.; HUBBARD, A. S.; WONG, E. F.; SCHALET, B. D.; HAVILAND, M. G.; KIMERLING, R. Response category functioning on the health care engagement measure using the nominal response model. **Assessment**, [S. l.], v. 30, n. 2, p. 375-389, 2023. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10262955/>. Acesso em: 07 out. 2023.

REVELLE, M. W. Procedures for psychological, psychometric, and personality. **R Package Biopsychn**, Northwestern University, Evanston, Illinois. R package version 2.3.6, 2023. Disponível em: URL: <https://CRAN.R-project.org/package=psych>. Acesso em: 07 out. 2023.

SILVA, A. F. Z. da; ANDRADE, D. F. de; BORGATTO, A. F.; NAKAMURA, L. R. Aplicação do modelo de reposta nominal da TRI a avaliação educacional de larga escala. **Sigmae**, Minas Gerais, v. 8, n. 2, p. 735-741, 2019. Disponível em: <http://publicacoes.unifal-mg.edu.br/revistas/index.php/sigmae/article/view/1036/693>. Acesso em: 07 out. 2023.

SMITH, T. I.; BENDJILALI, N. Motivations for using the item response theory nominal response model to rank responses to multiple-choice items. **Physical Review Physics Education Research**, [S. l.], v. 18, n. 1, p. 1-13, 2022. Disponível em: <https://journals.aps.org/prper/pdf/10.1103/PhysRevPhysEducRes.18.010133>. Acesso em: 07 out. 2023.

STEWART, J.; DRURY, B.; WELLS, J.; ADAIR, A.; HENDERSON, R.; MA, Y.; LEMONCHE, A. P.; PRITCHARD, D. Examining the relation of correct knowledge and misconceptions using the nominal response model. **Physical Review Physics Education Research**, [S. l.], v. 17, n. 1, p. 1-15, 2021. Disponível em: <https://journals.aps.org/prper/pdf/10.1103/PhysRevPhysEducRes.17.010122>. Acesso em: 07 out. 2023.

THISSEN, D.; STEINBERG, L.; FITZPATRICK, A. R. Multiple-Choice Models: the distractors are also part of the item. **Journal of Educational Measurement**, EUA, v. 26, n. 2, p. 161-76, 1989. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1745-3984.1989.tb00326.x>. Acesso em: 20 out. 2023.

VIALI, L.; CURY, H. N. Análise de erros em probabilidade: uma pesquisa com professores em formação continuada. **Educação Matemática Pesquisa**, [S. l.], v. 11, n. 2, p. 373-391, 2009. Disponível em: <http://funes.uniandes.edu.co/24348/1/Viali2009An%C3%A1lise.pdf>. Acesso em: 07 out. 2023.

VIANNA, H. M. Fundamentos de um programa de avaliação educacional. **Revista Meta: Avaliação**, Rio de Janeiro, v. 1, n. 1, p. 11-27, abr. 2009. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/11/4>. Acesso em: 29 ago. 2023.

ZHANG, X.; ZHAO, C.; XU, Y.; LIU, S.; WU, Z. Kernel causality among teacher self-efficacy, job satisfaction, school climate, and workplace well-being and stress. **TALIS. Frontiers in Psychology**, EUA, v. 12, p. 1-16, 2021. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.694961/full>. Acesso em: 07 out. 2023.

Contribuição dos(as) autores(as)

Sandra Cristina Martini Rostirola – Coleta e análise dos dados e elaboração do texto.

Elisa Henning – Análise dos resultados, orientação e elaboração do texto.

Ivanete Zuchi Siple – Análise dos resultados, orientação e elaboração do texto.

Correção de língua portuguesa: Marcia Vidal.

E-mail: frozzamarciavidal@gmail.com

Tradução: Diogo Guedes de Figueiredo

E-mail: diogoguedesdefigueiredo@gmail.com