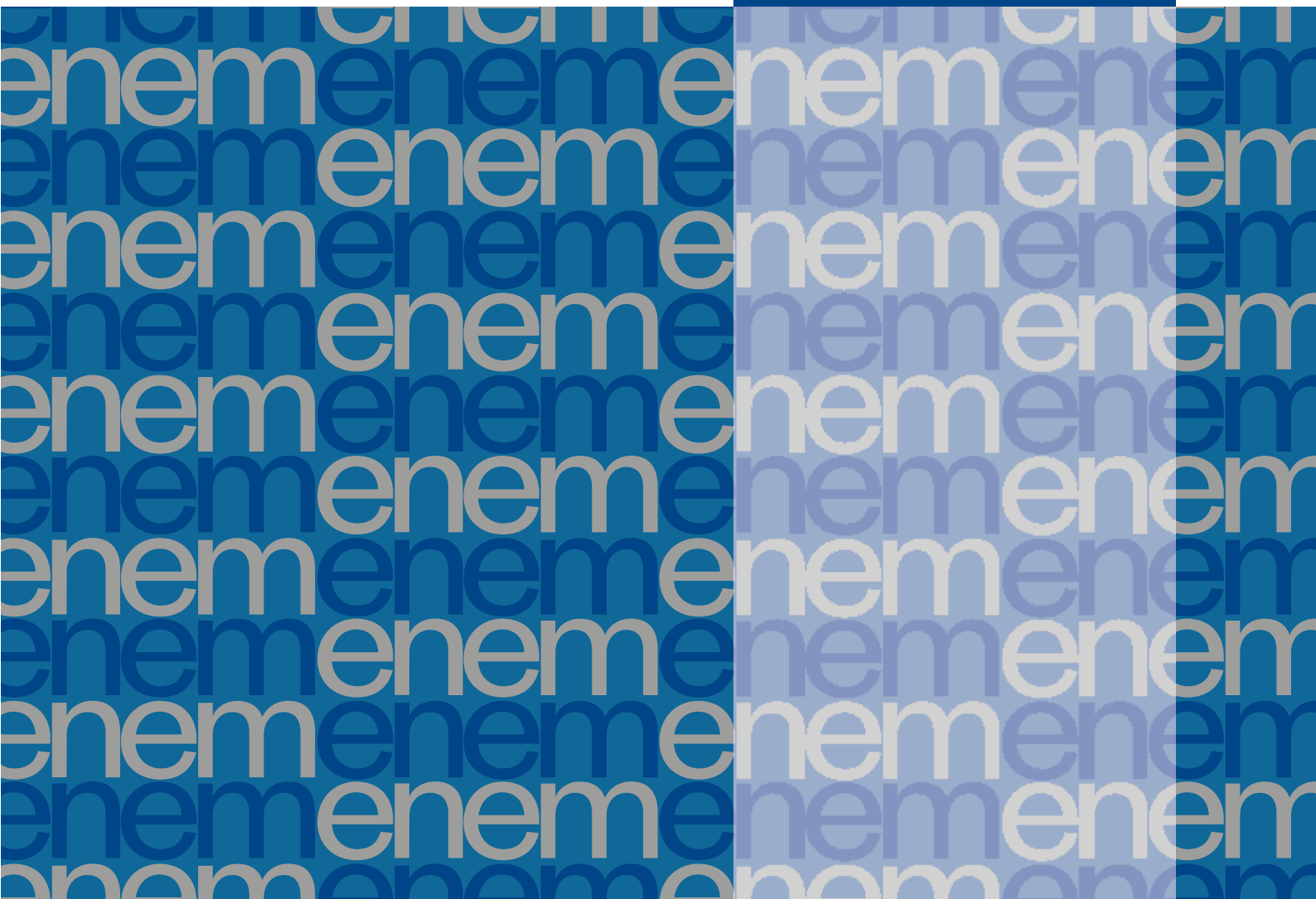


enem2021

**ENTENDA SUA NOTA
NO ENEM**

GUIA DO PARTICIPANTE

DIRETORIA DE AVALIAÇÃO
DA EDUCAÇÃO BÁSICA
DAEB



REPÚBLICA FEDERATIVA DO BRASIL

MINISTÉRIO DA EDUCAÇÃO | **MEC**

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS
EDUCACIONAIS ANÍSIO TEIXEIRA | **INEP**


DIRETORIA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA | **DAEB**



ENTENDA A SUA NOTA NO ENEM

GUIA DO PARTICIPANTE

Brasília-DF
Inep/MEC
2021



DIRETORIA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA (DAEB)

COORDENAÇÃO-GERAL DE INSTRUMENTOS E MEDIDAS (CGIM)

André Luiz Santos de Oliveira
Hemilhana Tolentina Pereira
Julio Maria do Carmo Neto
Melissa Castro de Oliveira Domingues
Natália Caixeta Barroso

Equipe técnica da 1ª edição do Entenda a sua nota no Enem: guia do participante.

Camila Akemi Karino
Eduardo Carvalho Sousa
Adriano Ferreti Borgatto – Universidade Federal De Santa Catarina (UFSC)
Dalton Francisco de Andrade – Universidade Federal De Santa Catarina (UFSC)

DIRETORIA DE ESTUDOS EDUCACIONAIS (DIRED)

COORDENAÇÃO DE EDITORAÇÃO E PUBLICAÇÕES (COEP)

PROJETO GRÁFICO/CAPA
Marcos Hartwich

PROJETO GRÁFICO/MIOLO
Raphael C. Freitas

DIAGRAMAÇÃO E ARTE-FINAL
Raphael C. Freitas

REVISÃO LINGUÍSTICA
Jair Santana Moraes

REVISÃO GRÁFICA
José Miguel dos Santos

ILUSTRAÇÕES
Gabriel Mário (Ascom)

Esta publicação deverá ser citada da seguinte forma:

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Entenda a sua nota no Enem: guia do participante. Brasília, DF: INEP, 2021.



SUMÁRIO

ESTA PUBLICAÇÃO POSSUI SUMÁRIO INTERATIVO
PARA RETORNAR AO SUMÁRIO, CLIQUE NO NÚMERO
DA PÁGINA EM CADA SEÇÃO

INTRODUÇÃO.....	5
1 CALCULANDO AS NOTAS.....	7
1.1 Calculando sua nota por meio do modelo de TRI	8
1.2 Posicionando sua nota na escala (régua)	9
2 ENTENDENDO OUTROS ASPECTOS DO MODELO DE TRI	13
2.1 Relação entre a nota da TRI e o número de acertos	13
2.2 Máximos e mínimos das provas.....	14
2.3 Composição das notas.....	15
2.4 Coerência das respostas	15
3 PROCEDIMENTOS DE CORREÇÃO	18
4 PROCESSO DE CRIAÇÃO DAS QUESTÕES	20
5 PRINCIPAIS DÚVIDAS SOBRE A NOTA DO ENEM.....	22
6 ENTENDENDO TECNICAMENTE A NOTA DA TRI	24
6.1 Estimação dos parâmetros dos itens	25
6.2 Estimação das proficiências (nota)	28
REFERÊNCIAS BIBLIOGRÁFICAS.....	29



INTRODUÇÃO

Como é calculada minha nota no Enem? Por que acertei mais questões este ano do que no ano anterior e a minha nota é menor? Qual é o peso de cada questão do Enem? Por que quem acerta todas as questões não recebe nota 1.000 ou os que erram todas não têm nota 0?

Para ajudar-lhe, este guia foi elaborado com o objetivo de esclarecer suas dúvidas. O Enem é uma prova em que a nota não é calculada da mesma forma como estamos acostumados nas salas de aula. Por isso, muitos participantes têm dificuldade para entendê-la.

Nas avaliações, normalmente, as nossas notas são calculadas de “0 a 10” ou de “0 a 100”, e a forma de calcular a nota é simples, bastando somar as questões corretas na prova. Imaginemos que um professor elabore uma prova com 10 questões, para avaliar o conhecimento dos seus alunos de Matemática, em que cada questão vale um ponto. Ao final da prova, o Aluno A e o Aluno B acertaram seis questões, porém não acertaram exatamente as mesmas questões. Será que os alunos deveriam receber a mesma nota? Será que os dois alunos possuem o mesmo conhecimento de Matemática?

Assim, a natureza específica de cada questão deve ser considerada na metodologia de cálculo, de tal forma que se possa diferenciar alunos com a mesma quantidade de acertos. Por exemplo, certas questões são acertadas consistentemente apenas por participantes com elevada habilidade, enquanto outras são frequentemente acertadas pela maioria. O modelo de Teoria de Resposta ao Item (TRI) utilizado no Enem capta isso. Neste texto, explicaremos o cálculo da nota da prova objetiva, uma vez que a nota da redação é calculada de outra forma, já explicada no documento *A redação no Enem 2020 – Cartilha do Participante*, disponível no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

Na prova objetiva do Enem, a nota não é calculada levando-se em conta somente o número de questões corretas, mas também a coerência das respostas do participante ao conjunto das questões que formam a prova. Essa coerência será discutida mais adiante. A nota é atribuída

em uma métrica (escala) criada pelo Inep especialmente para o Enem, com o objetivo de medir o conhecimento (proficiência) do participante em quatro áreas: Matemática e suas Tecnologias; Linguagens, Códigos e suas Tecnologias; Ciências da Natureza e suas Tecnologias; e Ciências Humanas e suas Tecnologias.

No Enem, para medir o conhecimento do participante, a prova é elaborada com 45 questões em cada área do conhecimento, segundo a matriz de referência do Exame, que toma como base os documentos oficiais que subsidiam o ensino médio. A proficiência é verificada com apoio na análise do perfil das respostas do participante ao conjunto de itens da prova.

NOTA: Neste guia item e questão serão tratados como sinônimos.

1 CALCULANDO AS NOTAS

Medir é avaliar algo por meio de um instrumento de medida e comparar com um padrão. Para medir distância, por exemplo, o instrumento é a régua, e o padrão pode ser o centímetro; para aferir a massa, o instrumento é a balança, e o padrão pode ser o grama. Já para medir o conhecimento das habilidades descritas nas matrizes do Enem, que é algo abstrato, o instrumento de medida é a prova, e o padrão dessa medida é a escala, construída por meio do modelo de TRI e da interpretação do construto pedagógico.



FIGURA 1

DESCRIÇÃO DO MODO DE AFERIÇÃO DO CONHECIMENTO

Fonte: Elaborado pela Ascom/Inep.
Nota: A proficiência é a nota do participante.

1.1 CALCULANDO SUA NOTA POR MEIO DO MODELO DE TRI

Os procedimentos de cálculo das notas dos participantes e de análise dos itens no Enem têm como base a Teoria de Resposta ao Item (TRI). A TRI é um conjunto de modelos matemáticos que busca representar a relação entre a probabilidade de o participante responder corretamente a uma questão, seu conhecimento na área em que está sendo avaliado e as características (parâmetros) dos itens. Um aspecto importante da TRI é que ela considera a particularidade de cada item. Assim, as notas não dependem do total de itens da prova, mas de cada item que a compõe. Dessa forma, duas pessoas com a mesma quantidade de acertos na prova são avaliadas de forma distintas a depender de quais itens estão certos e errados e podem, assim, ter notas diferentes.

O modelo matemático da TRI usado no Enem (ver detalhes técnicos adiante) considera três parâmetros que expressam as informações do item, essenciais para avaliar suas características e, conseqüentemente, a medida do conhecimento:

- parâmetro de discriminação*: é o poder de discriminação que cada questão possui para diferenciar os participantes que dominam dos participantes que não dominam a habilidade avaliada naquela questão (item);
- parâmetro de dificuldade*: associado à dificuldade da habilidade avaliada na questão, quanto maior seu valor, mais difícil é a questão. Ele é expresso na mesma escala da proficiência. Em uma prova de qualidade, devemos ter questões de diferentes níveis de dificuldade para avaliar adequadamente os participantes em todos os níveis de conhecimento;
- parâmetro de acerto casual*: em provas de múltipla escolha, um participante que não domina a habilidade avaliada em uma determinada questão da prova pode responder corretamente a um item devido ao acerto casual. Assim, esse parâmetro representa a probabilidade de um participante acertar a questão não dominando a habilidade exigida.

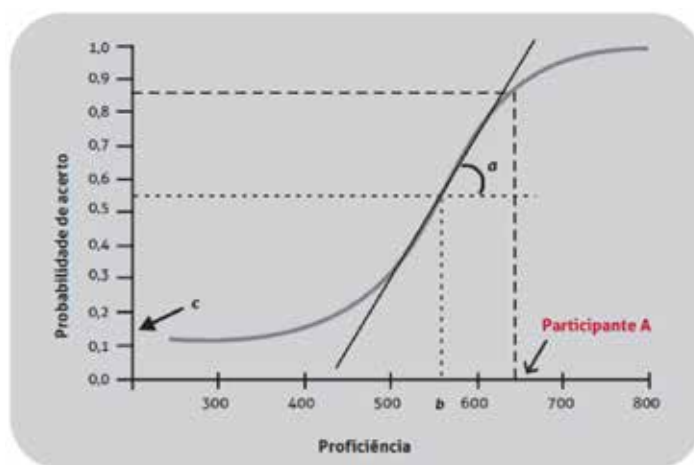


GRÁFICO 1

CURVA CARACTERÍSTICA DO ITEM

Fonte: Brasil. Inep ([2012], p. 13).

O Gráfico 1 apresenta o perfil de uma questão em relação a seus parâmetros. A curva é obtida pelo modelo de TRI e mostra a relação entre a probabilidade de acertos e a proficiência dos respondentes. Podemos notar, então, que somente respondentes com proficiência acima do valor do parâmetro b de dificuldade é que terão alta probabilidade de responder corretamente à questão. Um *Participante A*, com proficiência em torno de 650, tem, aproximadamente, probabilidade de 0,85 de responder corretamente à questão representada. Em outras palavras, espera-se que 85% dos participantes com proficiência 650 acertem essa questão ou, ainda, que a habilidade avaliada nela seja de domínio da maioria das pessoas que têm proficiência 650 ou mais.

De posse dos parâmetros dos itens é possível obter a nota do participante de modo objetivo, e participantes com exatamente o mesmo padrão de respostas apresentam exatamente as mesmas notas. O método utilizado para o cálculo da nota é o *Expected a Posteriori* (EAP), do qual há uma breve explicação na seção 6.2 deste documento.

1.2 POSICIONANDO SUA NOTA NA ESCALA (RÉGUA)

O Inep, a partir de pré-testagens nacionais e de algumas aplicações iniciais do Enem, obteve os valores dos parâmetros a , b e c de milhares de questões, montando assim um banco de itens (questões) para cada uma das quatro áreas. Com base nessas informações, as questões são posicionadas na escala (régua), permitindo, desse modo, sua interpretação pedagógica. O posicionamento de cada questão nos níveis da escala se dá por meio de critérios probabilísticos, os quais garantem que somente participantes com proficiência igual ou maior que a daquele nível possuem alta probabilidade de responder corretamente às questões que estão nesse nível e em níveis inferiores. Geralmente, o posicionamento do item está um pouco acima do parâmetro de dificuldade. A questão é sempre posicionada no valor da régua em que a probabilidade de acerto está próxima de 0,65, o qual representa que participantes nesse nível possuem alta probabilidade de dominar o conteúdo da questão.

Questões pedagogicamente mais fáceis serão posicionadas na parte inferior da régua e aquelas pedagogicamente mais difíceis serão posicionadas na parte superior. O grande diferencial da TRI é que questões e participantes são posicionados na mesma régua.

O Inep criou uma escala para cada área do conhecimento. Isso quer dizer que quatro notas são calculadas, sendo que cada escala depende de dois valores:

- valor de posição ou de referência, para o qual foi atribuído o valor 500, que representa o desempenho médio dos concluintes regulares do ensino médio de 2009 que realizaram o exame naquele ano;
- valor de dispersão, para o qual foi atribuído o valor 100, que representa uma medida de variabilidade média das notas desses concluintes em relação ao desempenho médio 500. Esse valor é conhecido como desvio padrão.

Conhecendo esses valores, podemos dizer que um participante com nota 600 apresenta proficiência com uma unidade de desvio padrão acima da proficiência média dos concluintes de 2009. Essa escala poderia ser representada graficamente por uma régua, como a apresentada a seguir, na qual as proficiências e os itens serão posicionados e seus pontos interpretados pedagogicamente.



FIGURA 2

ILUSTRAÇÃO DA ESCALA DE PROFICIÊNCIA DO ENEM

Fonte: Elaborado pela Ascom/Inep.

A escala tem uma interpretação pedagógica, e as questões são posicionadas na régua (escala) de acordo com uma coerência pedagógica. O participante também será posicionado na régua, ou seja, terá sua nota conforme seu conhecimento pedagógico. Isso quer dizer que sua nota não depende das notas dos demais participantes, mas apenas do posicionamento das questões na régua. A nota na escala (régua) pode assumir qualquer valor no conjunto dos números reais, como, por exemplo, 632,6, 421,8, etc.

No exemplo da figura, a seguir, é razoável dizer que o Participante A domina as habilidades avaliadas nas questões 1, 34, 45, 12 e 16, e não domina aquelas contempladas nas questões 19 e 35. Por sua vez, o Participante B deve dominar somente as habilidades exigidas nas questões 1 e 34. Logo, conclui-se que o Participante A deve dominar, além das habilidades contempladas nas questões 1 e 34, que são de domínio do Participante B, também aquelas das questões 12, 16 e 45, portanto sua nota deverá ser maior.

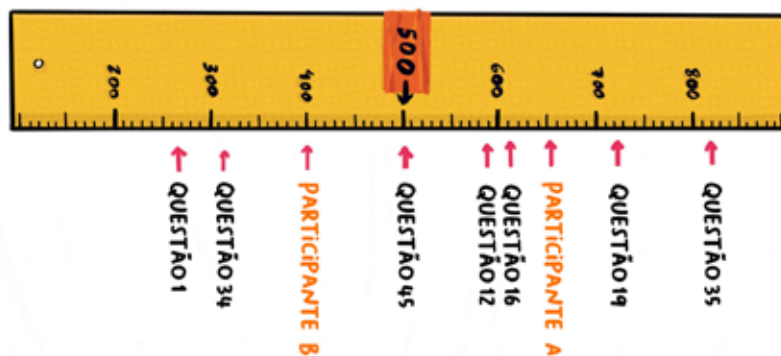


FIGURA 3

EXEMPLO DE ALOCAÇÃO DE QUESTÕES E PARTICIPANTES NA MESMA ESCALA

Fonte: Elaborado pela Ascom/Inep.

Para ilustrar a posição das questões na régua, o Inep desenvolveu um mapa de itens por área do conhecimento. Nesse mapa são apresentadas descrições de habilidades avaliadas em questões de provas. Abaixo são apresentadas algumas dessas descrições entre os níveis 650 e 700 de cada escala de proficiência. Você poderá visualizar o mapa completo e alguns itens no site do Inep.

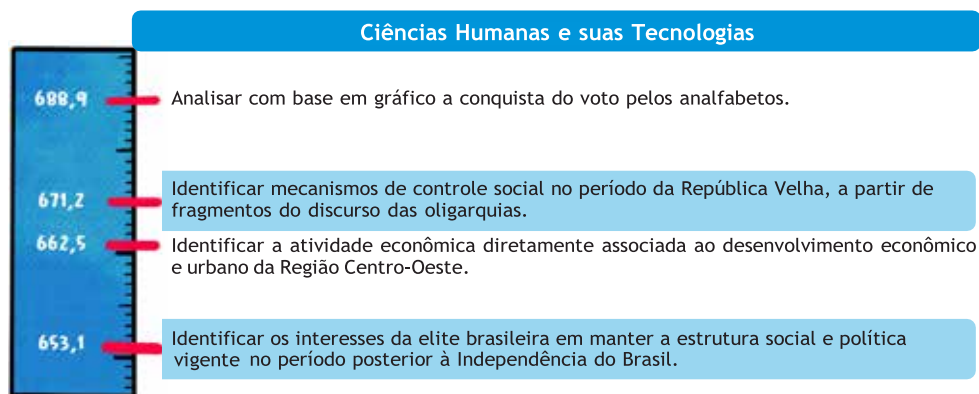


FIGURA 4

MAPA DE ITENS DE CIÊNCIAS DA NATUREZA E SUAS TECNOLOGIAS

Fonte: Elaborado pela Ascom/Inep.

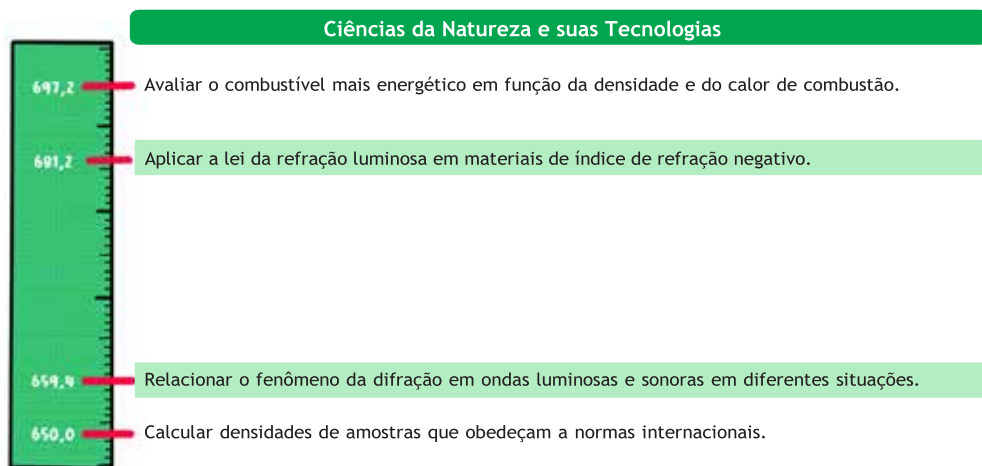


FIGURA 5

MAPA DE ITENS DE CIÊNCIAS HUMANAS E SUAS TECNOLOGIAS

Fonte: Elaborado pela Ascom/Inep.

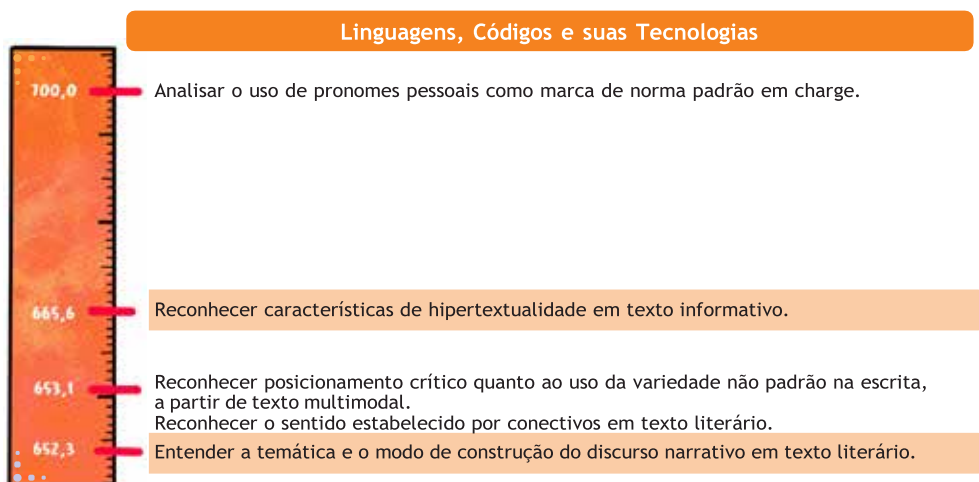


FIGURA 6

MAPA DE ITENS DE LINGUAGENS, CÓDIGOS E SUAS TECNOLOGIAS

Fonte: Elaborado pela Ascom/Inep.

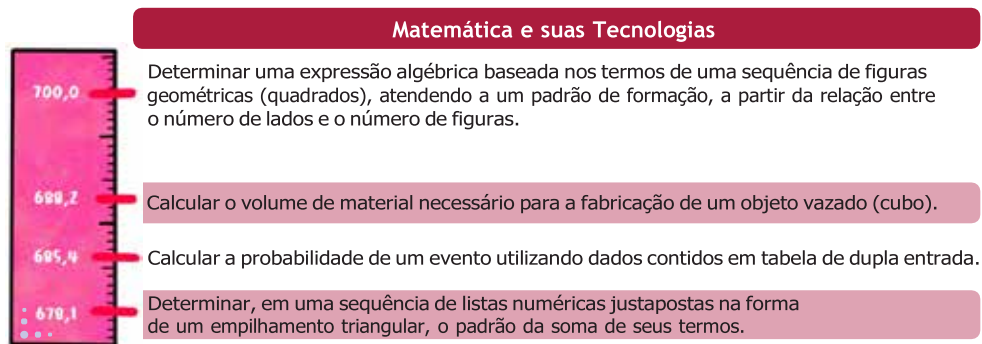


FIGURA 7

MAPA DE ITENS DE MATEMÁTICA E SUAS TECNOLOGIAS

Fonte: Elaborado pela Ascom/Inep.



2 ENTENDENDO OUTROS ASPECTOS DO MODELO DE TRI

Por se tratar de algo que usualmente não está presente no cotidiano de provas escolares, diversos aspectos da TRI se mostram como fontes de dúvida e preocupação por parcela dos estudantes brasileiros. Além disso, trata-se de uma metodologia estatística complexa, que não pode ser facilmente reproduzida sem o conhecimento técnico e uso de *softwares* específicos. Porém, é um método confiável e largamente utilizado em avaliações e exames educacionais em diversos países, como Estados Unidos, França, Holanda, Coreia do Sul e China, por exemplo; e em consolidadas avaliações como o Programa Internacional de Avaliação de Estudantes (Pisa) e o TOEFL, teste internacional de proficiência em língua inglesa.

Um exemplo de questionamento que surge com frequência se refere à relação das notas entre os participantes. Esclarecemos que a nota do participante depende somente de seu conhecimento e de seu momento na prova. Ela não depende do desempenho dos outros participantes, sejam eles da mesma edição ou de edições anteriores do exame. Também não possui relação com o desempenho médio dos concluintes regulares de 2009, pois este foi utilizado apenas como referência para a construção e interpretação pedagógica da escala na qual a proficiência (conhecimento) está representada.

Nesta seção iremos abordar aspectos teóricos da TRI, explanando pontos que usualmente geram confusão no entendimento do método, tais como: relação entre a nota e o número de acertos; máximos e mínimos das provas; composição das notas; coerências nas repostas.

2.1 RELAÇÃO ENTRE A NOTA DA TRI E O NÚMERO DE ACERTOS

Apesar de a nota do Enem não ser calculada diretamente pelo número de acertos, existe uma relação entre o número de acertos e a nota calculada pela TRI. Isso quer dizer que um

participante que teve um número de acertos alto terá nota alta no Enem, e um participante que teve pouco acerto terá nota baixa, notas essas relacionadas com os valores mínimo e máximo de cada prova. O gráfico, a seguir, mostra a relação entre a nota da TRI e o número de acertos de 300 participantes da prova do Enem 2020. Observe que se a nota da TRI fosse exatamente o número de acertos, os pontos nos gráficos estariam alinhados representando uma reta. As variações desses pontos em relação à reta mostram que participantes com o mesmo número de acertos podem ter notas diferentes no Enem. Note, no entanto, que essa variação não é tão grande.

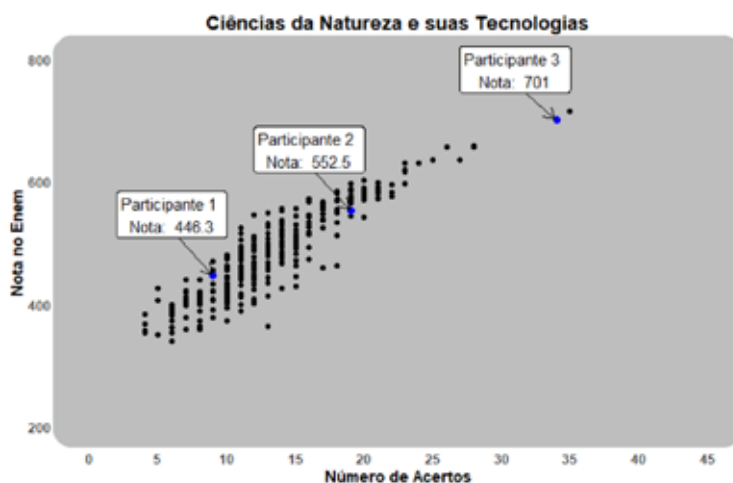


GRÁFICO 2

RELAÇÃO ENTRE O NÚMERO DE ACERTOS E A NOTA CALCULADA PELA TRI

Fonte: Elaborado pela Daeb/Inep.

O *Participante 1* teve nove acertos e sua nota foi 446,3; o *Participante 2* teve 19 acertos e sua nota foi 552,5 (lembre-se que a nota 500 representa a nota média dos concluintes regulares do ensino médio de 2009), ou seja, ele acertou menos da metade da prova, e seu desempenho foi equivalente ao desempenho médio dos concluintes regulares do ensino médio de 2009; e o *Participante 3* teve 34 acertos com nota 701,0 (a nota máxima, ou seja, para quem acertou 45 questões, foi 859,8).

Poderíamos nos perguntar: se existe alta relação entre o número de acertos e a nota da TRI, para que usar a nota da TRI? Não seria mais simples usar o número de acertos? Uma das vantagens de se utilizar a TRI é que sua nota nos possibilita fazer comparações entre notas de diferentes provas da mesma área, mas os acertos de provas diferentes não são diretamente comparáveis. Com ela, podemos também construir e interpretar pedagogicamente a escala de proficiência.

2.2 MÁXIMOS E MÍNIMOS DAS PROVAS

A prova é constituída por um conjunto de 45 itens com parâmetros, são esses parâmetros que definem os valores de mínimo e máximo da prova. Assim, quando a prova é composta com muitos itens fáceis, o máximo da prova tenderá a ser mais baixo e, quando ela é composta com muitos itens difíceis, o mínimo tenderá a ser mais alto.

Em outras palavras, o mínimo e o máximo em cada prova dependem somente das questões que a compõem e não de quem as responde. Apesar de as provas poderem apresentar mínimos e máximos diferentes, seus resultados são comparáveis, pois eles são todos calculados na mesma escala construída com base numa única matriz de competências.

Um aspecto do modelo de TRI utilizado no Enem é não ter um limite padrão inferior ou superior. Isso significa que as proficiências dos participantes não variam entre zero e mil. Os valores máximos e mínimos de cada prova dependerão das características dos itens selecionados.

2.3 COMPOSIÇÃO DAS NOTAS

Na divulgação dos resultados cada participante receberá cinco notas, uma para cada área da prova objetiva e a nota da redação. Cada faculdade/universidade tem regras próprias para utilizar as notas de ingresso, colocando peso nas notas das áreas que julgue ser mais importantes para o curso que está sendo pleiteado ou fazendo uma média aritmética simples.

2.4 COERÊNCIA DAS RESPOSTAS

No cálculo da nota, o modelo matemático de TRI usado no Enem considera a coerência das respostas corretas do participante. Espera-se que participantes que acertaram as questões difíceis devam também acertar as questões fáceis, pois, entende-se que a aquisição do conhecimento ocorre de forma cumulativa, de modo que habilidades mais complexas requerem o domínio de habilidades mais simples, lembrando que o posicionamento das questões na escala de proficiência é determinado pelas respostas dos estudantes. Assim, na régua de Ciências da Natureza, um participante que é capaz de responder a uma questão que exige a avaliação de qual é o combustível mais energético em função da densidade e do calor de combustão (posicionado no nível 697,2 da escala) tem alta probabilidade de responder a uma questão que exige o cálculo de densidade de amostras que obedeçam às normas internacionais, que está posicionada no nível 650 da escala; no entanto, se o participante acerta a questão de um nível acima, que utiliza conhecimentos de um nível abaixo, podemos inferir que não existe coerência nas respostas e, dessa forma, os acertos acrescentarão menos na nota do participante na régua.

Cabe ressaltar que a comparação entre os resultados das diferentes áreas deve ser feita em termos pedagógicos, ou seja, utilizando a interpretação pedagógica da escala de cada área. A comparação entre uma nota 800 em Matemática e suas Tecnologias e uma nota 750 em Ciências da Natureza e suas Tecnologias só deve ser feita em termos de quanto de conhecimento em Matemática e suas Tecnologias e em Ciências da Natureza e suas Tecnologias esses valores representam pedagogicamente.

A seguir, temos as respostas a 10 questões já posicionadas, dadas por dois participantes. Ambos acertaram a mesma quantidade de questões (5), mas o *Participante A* acertou as questões mais fáceis e, após certo nível de dificuldade, passou a errar, conforme o esperado pedagogicamente, logo sua nota será maior do que a do *Participante B*, pois este acertou as

questões mais difíceis e errou as mais fáceis. Entende-se, de acordo com o modelo da TRI utilizado no Enem, que foram acertadas “no chute”. Esse controle é devido ao parâmetro de acerto casual.

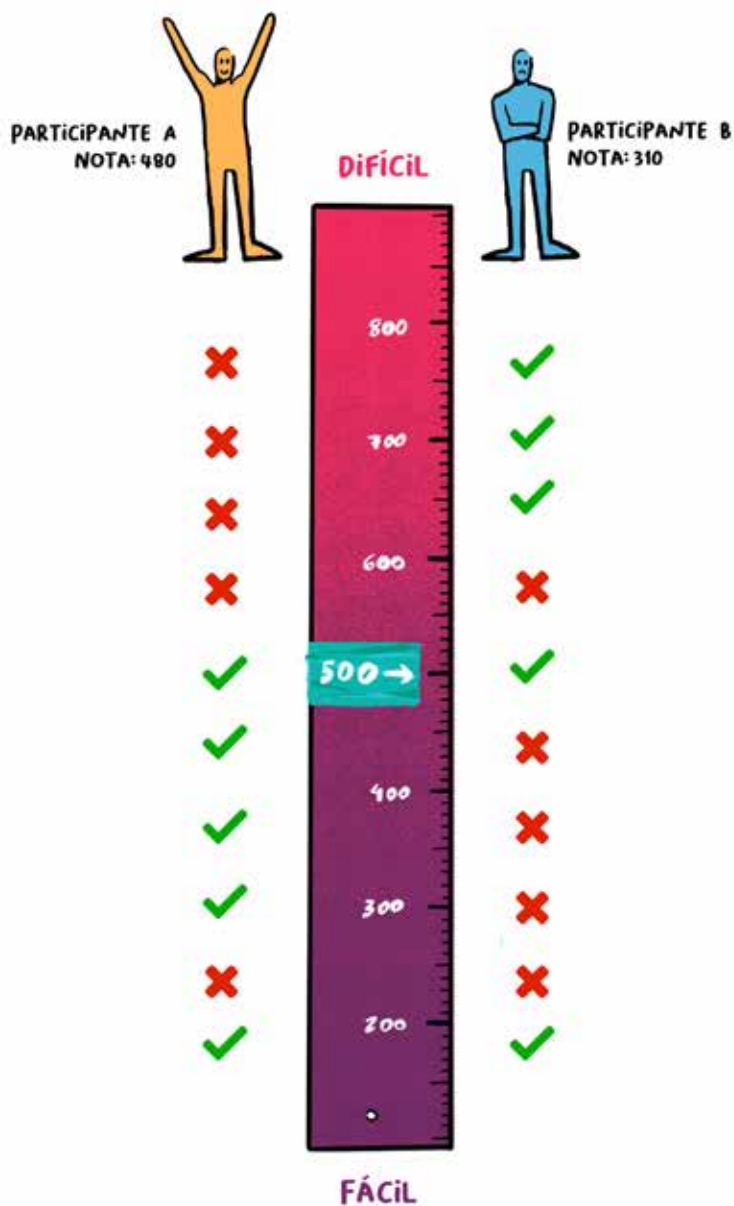


FIGURA 8

PADRÃO DE RESPOSTAS DE DOIS PARTICIPANTES A 10 QUESTÕES E SUAS RESPECTIVAS PROFICIÊNCIAS

Fonte: Elaborado pela Ascom/Inep.

Nota: Quando dizemos que o participante acertou uma questão “no chute”, não significa que sua nota irá diminuir, mas ela não tem tanto valor como se o participante tivesse acertado os itens com a coerência pedagógica esperada. Então, sempre é melhor responder à questão do que a deixar em branco, pois uma questão certa sempre aumenta a nota, e uma questão deixada em branco é corrigida como errada.

A ilustração a seguir também exemplifica a coerência esperada.

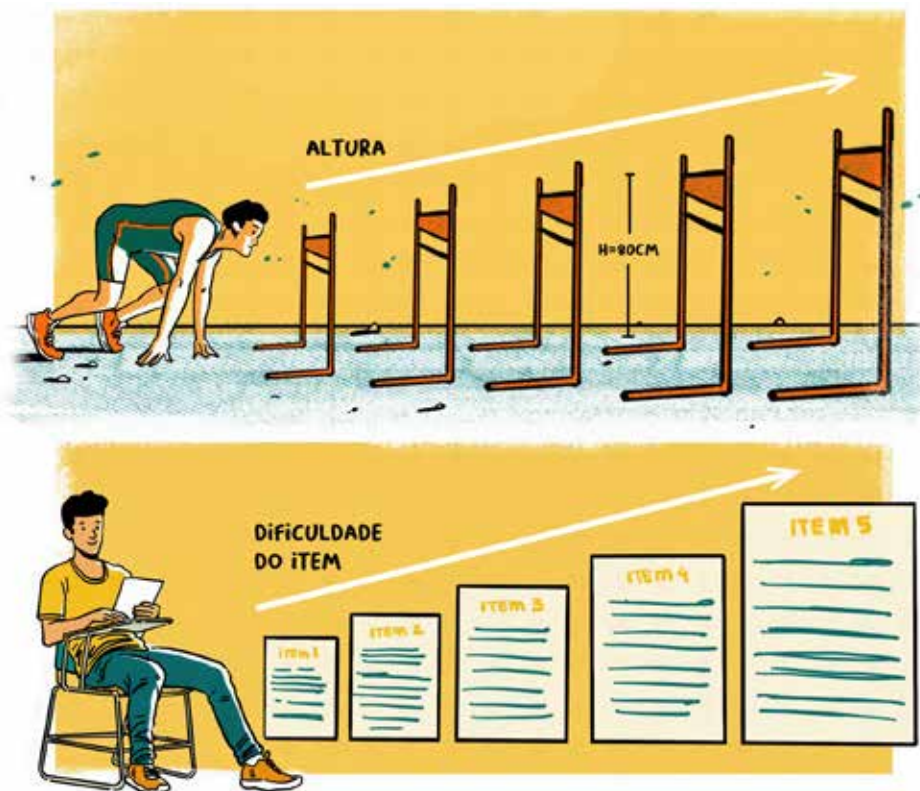


FIGURA 9
EXEMPLIFICAÇÃO DE COERÊNCIA PEDAGÓGICA

Fonte: Elaborado pela Ascom/Inep.



3 PROCEDIMENTOS DE CORREÇÃO

Considerando que o cálculo das notas de acordo com a TRI exige um conhecimento avançado de estatística e a utilização de recurso computacional, os cálculos são realizados de forma independente por grupos distintos de especialistas, todos profissionais com larga experiência na área e com formação em estatística, matemática e/ou psicometria. Esse procedimento de múltiplas conferências garante a qualidade dos resultados do Enem. O cálculo da nota exige 6 etapas, apresentadas no infográfico a seguir:



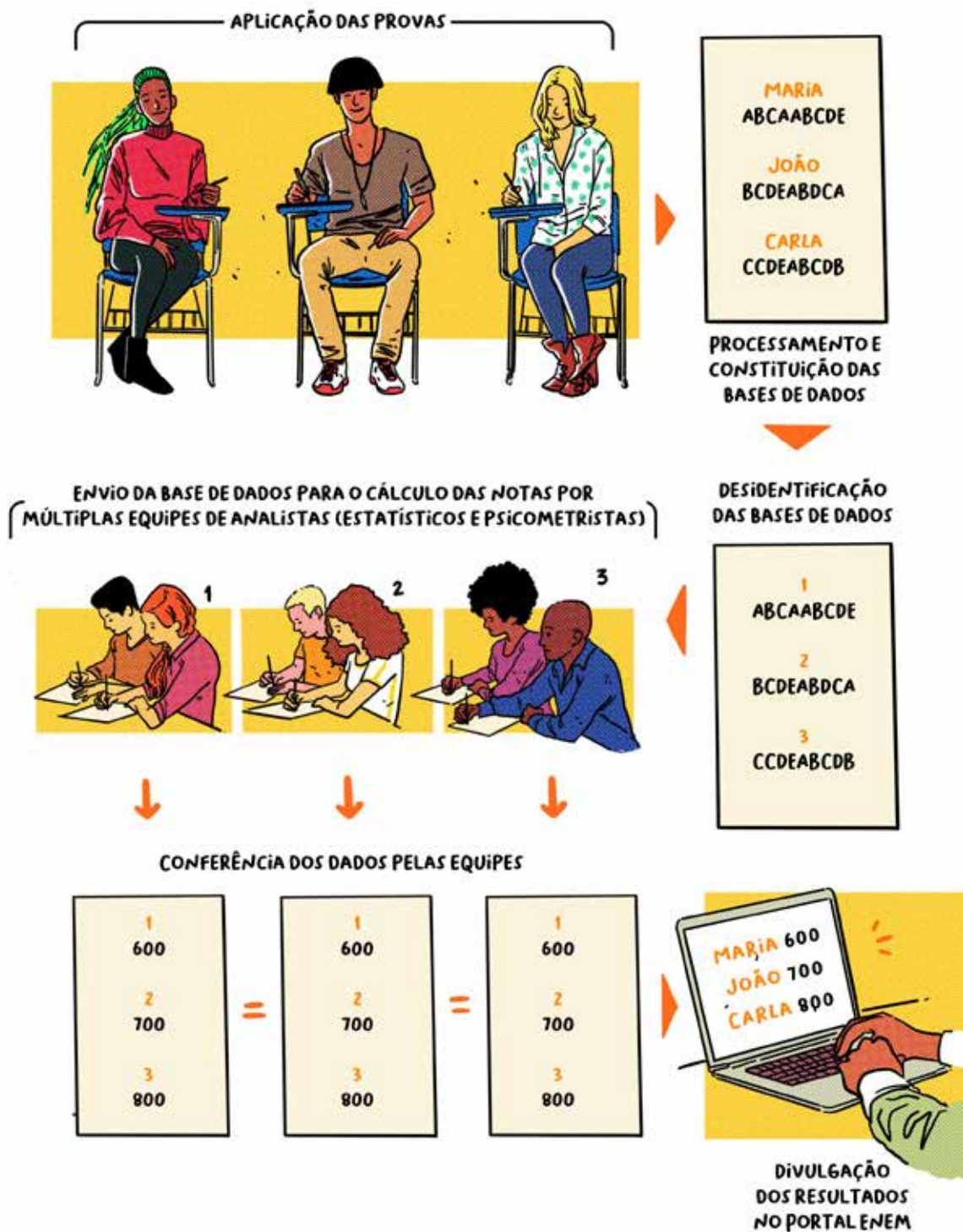


FIGURA 10

PROCESSO DE ANÁLISE E DIVULGAÇÃO DOS RESULTADOS DO ENEM

Fonte: Elaborado pela Ascom/Inep para este guia



4 PROCESSO DE CRIAÇÃO DAS QUESTÕES

O Enem exige um trabalho contínuo de criação de itens para que as provas tenham itens válidos e de boa qualidade. Com o objetivo de manter a organização dos itens, foi criado o Banco Nacional Itens (BNI).

O BNI é uma base de dados que tem como principal objetivo atender às necessidades do Inep na construção das provas, em especial, a do Enem. A construção desse banco envolve algumas etapas essenciais para o desenvolvimento de uma ferramenta confiável e válida:

- *Elaboração de itens* que atenda à matriz de referência. O Inep realiza chamadas públicas para a construção de uma rede de educadores e pesquisadores da educação brasileira com experiência docente, interessados em elaborar e revisar itens para o BNI, constituindo o Banco de Colaboradores do BNI. Os elaboradores devem ser especialistas nos conteúdos, assim como nos critérios técnicos;
- *Validação pedagógica dos itens* em que uma equipe formada por especialistas revisa os itens checando a sua adequação com a matriz de referência, conteúdo, características técnicas, gabarito e uso de linguagem apropriada.
- *Pré-testagem dos itens* em uma população similar à população que fará a prova do Enem. Após essa pré-testagem, os itens são analisados estatisticamente e os resultados examinados para que os itens apropriados sejam identificados no BNI para utilização em uma prova. Por meio dessa análise, é possível verificar os itens com desempenho apropriado e os itens que devem ser reelaborados ou eliminados das futuras versões da prova. O objetivo central da pré-testagem é a melhoria dos itens e, conseqüentemente, a melhoria dos instrumentos e das inferências.



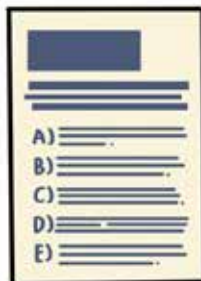
ELABORAÇÃO DOS ITENS
O INEP REALIZA PARCERIAS COM PROFESSORES CREDENCIADOS EM CHAMADAS PÚBLICAS OU INSTITUIÇÕES DE EDUCAÇÃO SUPERIOR CREDENCIADAS INTERESSADAS EM ELABORAR E REVISAR ITENS PARA A COMPOSIÇÃO DAS PROVAS DO ENEM. O OBJETIVO DESSAS PARCERIAS É AUMENTAR A PARTICIPAÇÃO DA COMUNIDADE ACADÊMICA DE TODO O BRASIL NOS PROCESSOS DE AVALIAÇÃO EDUCACIONAL, AGREGANDO EXPERIÊNCIA, CONHECIMENTO E DIVERSIDADE.



PRÉ-TESTE

O PRÉ-TESTE CONSISTE NA APLICAÇÃO DE UM CONJUNTO DE ITENS A UMA AMOSTRA DE ALUNOS COM CARACTERÍSTICAS SEMELHANTES ÀS DA POPULAÇÃO PARA A QUAL A PROVA SE DESTINA. É A FORMA EMPÍRICA DE SE AVALIAR A QUALIDADE TÉCNICO-PEDAGÓGICA E PSICOMÉTRICA DOS ITENS. ESSA ETAPA TEM COMO OBJETIVO CAPTAR SUBSÍDIOS IMPORTANTES PARA AUMENTAR A PRECISÃO DA PROVA QUE SERÁ APLICADA A MILHÕES DE PARTICIPANTES DO ENEM.

BANCO NACIONAL DE ITENS (BNI)
É GERENCIADO PELO INEP E CONSISTE EM UM SISTEMA DE BANCO DE DADOS NO QUAL OS ITENS DE PROVAS SÃO ARMAZENADOS DE FORMA SEGURA. ESSES ITENS FICAM DISPONÍVEIS PARA A CONSTRUÇÃO DE PROVAS. A MANUTENÇÃO DO BNI DEPENDE DA ENTRADA CONSTANTE DE ITENS DE QUALIDADE.



ANÁLISE DEPOIS DO PRÉ-TESTE
A PARTIR DAS RESPOSTAS DOS ALUNOS, REALIZA-SE UMA SÉRIE DE ANÁLISES ESTATÍSTICAS E PEDAGÓGICAS; POR EXEMPLO, AVALIAM-SE A DIFICULDADE DO ITEM, A CAPACIDADE DE DISCRIMINAÇÃO E A POSSIBILIDADE DE ACERTO AO ACASO. DEPOIS DESSAS ANÁLISES, AS QUESTÕES QUE ATENDEM AOS CRITÉRIOS FICAM DISPONÍVEIS PARA A MONTAGEM DAS PROVAS, E AS DEMAIS QUESTÕES SÃO DESCARTADAS OU ENCAMINHADAS PARA REFORMULAÇÃO.

ELABORAÇÃO DAS PROVAS
DURANTE A SELEÇÃO DOS ITENS PARA A COMPOSIÇÃO DE UMA PROVA, OS ÍNDICES PSICOMÉTRICOS OBTIDOS A PARTIR DO PRÉ-TESTE SÃO UTILIZADOS. OUTRAS CARACTERÍSTICAS CONSIDERADAS SÃO: CONTEÚDO ABORDADO, TEMÁTICA E HABILIDADE DA MATRIZ DE REFERÊNCIA.



QUESTÃO ADEQUADA
ENCAMINHADA AO BANCO NACIONAL DE ITENS E UTILIZADA PARA A ELABORAÇÃO DAS PROVAS.

QUESTÃO INADEQUADA
RETORNA PARA REFORMULAÇÃO OU DESCARTE.



FIGURA 11

PROCESSO DE CONSTRUÇÃO DA PROVA DO ENEM

Fonte: Elaborado pela Ascom/Inep.



5 PRINCIPAIS DÚVIDAS SOBRE A NOTA DO ENEM

Responderemos, a seguir, a algumas questões mais frequentes, perguntadas pelos participantes nas redes sociais.

Por que tirei uma nota menor neste ano, se acertei um número maior de questões que no ano passado?

A estimativa da proficiência na TRI não é calculada diretamente pelo número de acertos; ela envolve os parâmetros dos itens que foram tanto acertados, quanto errados. Dessa forma, a causa desse resultado pode decorrer tanto da última prova ser composta por itens mais fáceis como das características do seu padrão de resposta nela, que poderia ser mais coerente com uma proficiência menor, por exemplo, caso você tenha errado itens fáceis.

Como calcular a média das minhas notas no Enem?

Cabe ao Inep apenas divulgar as notas por área do conhecimento e da redação. Isso significa que cada Instituição de Ensino Superior (IES) tem autonomia para definir os critérios de uso das notas do Enem para ingresso na educação superior.

Por que minha nota é só 706,7 se eu acertei 40 questões na prova de Ciências Humanas em 2020?

Apesar do número grande de acertos, é preciso analisar a coerência das respostas na prova. Provavelmente, das cinco questões que você errou, deve haver algumas questões bem fáceis e, por isso, o acréscimo da sua nota não foi maior. É importante lembrar também que a nota máxima na prova de Ciências Humanas e suas Tecnologias, na edição de 2020, foi 862,6, ou seja, seria essa sua nota caso você tivesse acertado todas as questões.

Se acertei mais questões em Ciências Humanas que em Matemática, por que minha nota em Matemática é maior?

Cada área de conhecimento possui sua própria escala, logo as notas não são comparáveis diretamente entre as áreas. O importante é você verificar o significado pedagógico da nota através da escala pedagógica construída pelo Inep para cada uma das áreas.

Qual é o peso de cada questão do Enem utilizado no cálculo da nota?

No Enem, não são utilizados pesos para o cálculo das notas. A estimação da proficiência tem como base a consistência das respostas e as características (parâmetros) de cada questão. Entretanto, esses parâmetros não podem ser interpretados como pesos, e está errada a inferência de que a divulgação dos parâmetros possibilitaria a reprodução dos cálculos com simples ponderações.

A utilização da TRI no cálculo dos resultados torna a prova do Enem mais difícil?

O uso da TRI no cálculo dos resultados não altera a dificuldade da prova. A TRI não modifica, significativamente, a classificação dos participantes em relação ao percentual de acertos. Contribui, sim, para detalhar melhor as notas, o que ajuda a evitar grande número de empates. Aplicando ou não a TRI, o mais preparado se sairá melhor na prova.

Como sei se fui bem na prova depois da divulgação das notas?

Meses depois da aplicação do exame, o Inep divulga em seu site, na página de acompanhamento do Exame, o boletim do participante, o qual registra a porcentagem de participantes do Enem por faixa de proficiência (abaixo de 300, 300 a 400, 400 a 500, 500 a 600 e, assim, sucessivamente). Com base nessa distribuição, é possível verificar em qual faixa sua nota se posiciona.

Quais as notas mínima e máxima nas provas objetivas?

As notas mínimas e máximas variam e dependem das questões da prova. Como as questões das provas não são as mesmas, em cada ano, temos notas mínima e máxima diferentes. Na divulgação dos resultados, o Inep disponibilizará as notas mínima e máxima das provas objetivas.

6 ENTENDENDO TECNICAMENTE A NOTA DA TRI

Antes de começar a explicar o cálculo da nota pela TRI, iremos comentar a propriedade de invariância dos parâmetros. Essa propriedade indica que os valores dos parâmetros dos itens não dependem da habilidade de quem os responde e a proficiência (nota) do respondente é estimada independentemente do conjunto de itens utilizado no teste. Essa propriedade tem grande relevância no Enem, pois garante que o participante receba notas estatisticamente iguais em qualquer uma das versões do exame que ele faça.

A TRI modela a probabilidade de um participante responder corretamente a um item (questão), em função dos parâmetros deste e da proficiência do participante. Essa relação é expressa por meio de uma função monotônica crescente, a qual indica que quanto maior o conhecimento do participante, maior será sua probabilidade de acertar o item (ver, por exemplo, Andrade; Tavares; Valle, 2000; Baker; Kim, 2004; Hambleton; Swaminathan; Rogers, 1991; Klein, 2003; Pasquali, 1997). No Enem, a função monotônica é uma função logística de três parâmetros, desenvolvida por Birnbaum (1968), em que a probabilidade P de uma resposta correta do participante j ao item i ($u_{ji} = 1$), em função do parâmetro de proficiência θ_j e dos parâmetros do item a_i , b_i e c_i , é dada por

$$P(x_j = 1 | \theta, a_i, b_i, c_i) = c_i + \frac{(1 - c_i)}{1 + \exp[a_i(\theta - b_i)]} = P_{j1}(\theta)$$

Onde:

x_j é a resposta ao item i . $x_j = 1$ se a resposta estiver correta e $x_j = 0$, caso contrário;

θ_j representa a habilidade (traço latente) do participante j ;

a_i é o parâmetro de inclinação do item i , também chamado de parâmetro de discriminação do item, o qual é sempre maior que zero;

b_i é o parâmetro de posição do item, também chamado de parâmetro de dificuldade; e c_i é o parâmetro da assíntota inferior do item i , refletindo as chances de um estudante de proficiência muito baixa selecionar a opção de resposta correta. Esse parâmetro recebe um valor entre 0 e 1.

Denota-se a probabilidade de resposta errada ao item de:

$$P_{j0}(\theta) = P(x_j = 0 | \theta, a_i, b_i, c_i) = 1 - P_{j1}(\theta)$$

O modelo logístico de três parâmetros utilizado no Enem pressupõe a unidimensionalidade do teste, ou seja, que todos os itens apresentados em uma área de conhecimento do Exame devem medir um único traço latente. Apesar de sabermos que qualquer tarefa que realizamos envolve mais de uma habilidade na sua execução, para a TRI é suficiente supor que haja uma habilidade dominante em cada área de conhecimento, e é esta que se supõe estar sendo medida no teste.

Outro pressuposto deste modelo é o de independência local. Isso quer dizer que a resposta do participante a um item não pode influenciar sua resposta a outros itens. Esse pressuposto está diretamente ligado ao de unidimensionalidade. Se todos os itens medem uma única dimensão e a posição do participante nesta dimensão não se altera durante o exame, então as respostas aos itens são estatisticamente independentes. Essa suposição é muito importante para concluirmos que a sequência de respostas do participante a uma série de itens será o produto das probabilidades de cada item. Falaremos mais sobre isso posteriormente.

Os parâmetros dos itens são estimados conjuntamente para todos os itens de uma mesma área de conhecimento em um conjunto de testes, como, por exemplo, pré-testes e Enem, mediante o procedimento de estimação bayesiano marginal. A extensão da TRI de um grupo de participantes para vários é feita considerando-se distribuições *a priori* diferentes para cada um deles. Um grupo de referência é escolhido para receber uma distribuição *a priori* e, dessa maneira, eliminar a indeterminação do modelo. Resolvida essa questão, estimam-se os parâmetros das distribuições *a priori* dos outros grupos e os parâmetros dos itens.

6.1 ESTIMAÇÃO DOS PARÂMETROS DOS ITENS

O método de máxima verossimilhança marginal é utilizado para estimar os parâmetros dos itens, sem considerar as distribuições *a priori* deles, no caso de um grupo de participantes.

Seja ξ'_i o vetor de parâmetros do item i e $\xi' = (\xi'_1, \dots, \xi'_m)$ o vetor de parâmetros de todos os itens.

Seja $\chi' = (\chi_1, \dots, \chi_m)$ um vetor de respostas, $\chi_j = 1$ ou 0 , de um participante com proficiência θ . A probabilidade de ocorrência deste vetor de resposta é:

$$P(\chi' | \theta, \xi) = \prod_{j=1}^m P_{j1}^{\chi_j} (1 - P_{j1})^{(1-\chi_j)} = \prod_{j=1}^m P_{j1}^{\chi_j} P_{j0}^{(1-\chi_j)} = \prod_{j=1}^m P_{j1}^{\chi_j} P_{j0}^{1-\chi_j}$$

Onde:

$x_{j1}=1$ se $x_j=1$, e 0 caso contrário,
 $x_{j0}=1$ se $x_j=0$, e 0 caso contrário e
 m é o número de itens.

Supondo-se que os participantes pertençam a uma população em que a proficiência θ tenha uma distribuição de probabilidade contínua, com densidade $g(\theta)$ e média e variância finitas, tem-se que a probabilidade marginal é dada por:

$$P(x / \xi) = \int P(x / \theta, \xi) g(\theta) d\theta$$

e a função de verossimilhança marginal de ξ é dada por:

$$L = P(x_1, \dots, x_n / \xi) = \prod_{j=1}^n P(x_j / \xi)$$

Onde n é o número de participantes do teste.

Em geral, a integral $\int P(x / \theta, \xi) g(\theta) d\theta$ é estimada pela fórmula de quadratura gaussiana:

$$P(x / \xi) \approx \sum_{q=1}^Q P(x / X_q, \xi) A(X_q)$$

Onde:

X_q é um ponto de quadratura;

$A(X_q)$ é um peso positivo correspondente à função de densidade $g(\theta)$ no ponto X_q .

No Enem são utilizados 40 pontos de quadratura.

Maximizar L é o equivalente a maximizar $\log L$, que pode ser resolvida a partir do sistema a seguir:

$$\frac{\partial \log L}{\partial \xi_{i\ell}} = \sum_{q=1}^Q \sum_{k=0}^1 \tilde{r}_{ikq} \frac{1}{P_{jk}(X_q, \xi_i)} \frac{\partial P_{jk}(X_q, \xi_i)}{\partial \xi_{i\ell}}$$

$$i=1, \dots, m; \ell = 1, \dots, \ell_i$$

onde m é o número de itens, ℓ_i é o número de parâmetros do item i e \tilde{r}_{ikq} é o número esperado de respostas na categoria $k=0$ ou 1 , do item i por participantes com proficiência no intervalo $(X_q - \Delta X_q / 2, X_q + \Delta X_q / 2)$.

O pode ser estimado pela fórmula a seguir:

$$\tilde{r}_{ikq} = \sum_{j=1}^n x_{ijk} P(X_q, x_j, \xi) \Delta X_q$$

\tilde{N}_{ikq} é o número esperado de participantes que responderam ao item i com proficiência no intervalo $(X_q - \Delta X_q / 2, X_q + \Delta X_q / 2)$ e é dado por:

$$\tilde{N}_{ikq} = \sum_{k=0}^1 \tilde{r}_{ikq}$$

A probabilidade *a posteriori* da proficiência estar no intervalo $(X_q - \Delta X_q / 2, X_q + \Delta X_q / 2)$ dado o vetor de resposta x_j e o vetor de parâmetros ξ é dada por

$$P(X_q / x_j, \xi) \Delta X_q = \frac{P(x_j / X_q, \xi) A(X_q)}{\sum_{q=1}^Q P(x_j / X_q, \xi) A(X_q)}$$

Por meio do algoritmo E-M (Esperança-Maximização), obtém-se a solução do sistema

$$\frac{\partial \log L}{\partial \xi_{it}} = 0$$

$$i=1, \dots, m; \ell = 1, \dots, \ell_i$$

Esse algoritmo é executado em dois passos e funciona da seguinte maneira:

Passo E. A partir dos valores de ξ'_1, \dots, ξ'_m , calcula-se \tilde{r}_{ikq} e \tilde{N}_{ikq} .

Passo M. Dados \tilde{r}_{ikq} e \tilde{N}_{ikq} , encontra-se ξ que resolve o sistema.

Após cada passo M, comparam-se os novos estimadores de ξ com os anteriores, e o ciclo E-M é repetido até que os estimadores de ξ fiquem estáveis.

Supõe-se que a distribuição da proficiência segue uma distribuição normal de média 0 e variância 1.

A estimação dos parâmetros por esse método de máxima verossimilhança marginal pode apresentar problemas de convergência por conta do modelo logístico de três parâmetros utilizado no Enem. Por isso, utiliza-se o método bayesiano marginal.

Nesse método é considerada uma distribuição contínua *a priori* para os parâmetros ξ dos itens, dada pela densidade $g(\xi)$.

O estimador bayesiano é o valor de ξ que maximiza a densidade da distribuição *a posteriori*

$$g(\xi / x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n / \xi) g(\xi)}{P(x_1, \dots, x_n)}$$

Onde $P(x_1, \dots, x_n / \xi) = L$ é a função de verossimilhança marginal de ξ .

Como maximizar $g(\xi / x_1, \dots, x_n)$ é equivalente a maximizar $\log L + \log g(\xi)$, pode-se utilizar o procedimento descrito anteriormente apenas acrescentando as distribuições *a priori* dos parâmetros.

Quando se tem múltiplos grupos na análise, cada um deles possui uma distribuição contínua com densidade $g_k(\theta)$ e média e variância finitas. Em geral, supõe-se a distribuição normal para todos os grupos, mas com médias e variâncias diferentes. Para resolver a indeterminação do modelo, são utilizadas média 0 e variância 1 para o grupo de referência. A verossimilhança marginal neste caso é dada por:

$$L = \prod_{k=1}^K \prod_{i=1}^{n_k} P(x_{ki} / \xi, \eta_k) = \prod_{k=1}^K \prod_{i=1}^{n_k} \int P(x_{ki} / \theta, \xi) g_k(\theta) d\theta$$

onde k é o número de grupos, n_k o número de participantes no grupo e η_k é o vetor de parâmetros da distribuição do grupo k .

Os métodos de estimação dos parâmetros dos itens são os mesmos no caso de um grupo. O que muda no caso de múltiplos grupos é que as médias e variâncias das distribuições dos outros grupos precisam ser estimadas conjuntamente com os parâmetros dos itens.

6.2 ESTIMAÇÃO DAS PROFICIÊNCIAS (NOTA)

Para se obter a nota do participante do Enem, basta estimar (encontrar) o valor do parâmetro θ , que representa a nota do participante, levando em consideração o vetor de respostas dele e os parâmetros dos itens que compõem a prova. Como não é possível obter esse valor diretamente, mediante uma fórmula simples, é preciso utilizar um método matemático mais sofisticado para calcular a nota. Existem, na literatura, alguns métodos para o cálculo da nota, tais como a máxima verossimilhança, a moda da *posteriori* e a média da *posteriori* (EAP). O Enem e outras avaliações de larga escala, como, por exemplo, o Sistema de Avaliação da Educação Básica (Saeb), utilizam o método EAP.

Esse método tem como princípio usar uma função de probabilidade *a priori* para calcular a nota. Essa função tem a informação da nota média(0) e do desvio padrão(1) dos participantes que pertencem ao grupo de referência. Com base nessa informação, as notas são calculadas com mais precisão para a maior parte dos participantes, por meio da fórmula apresentada a seguir:

$$\hat{\theta} \approx \frac{\sum_{q=1}^Q X_q P(x_j / X_q, \xi) A(X_q)}{\sum_{q=1}^Q P(x_j / X_q, \xi) A(X_q)}$$

Sua precisão é medida com o desvio padrão *a posteriori* (*posterior standard deviation*, PSD) estimado por:

$$PSD(\hat{\theta}_j) \approx \frac{\sum_{q=1}^Q (X_q - \hat{\theta}_j)^2 P(x_j / X_q, \xi) A(X_q)}{\sum_{q=1}^Q P(x_j / X_q, \xi) A(X_q)}$$

Para aqueles que querem se aprofundar no assunto, recomendamos: Andrade, Tavares e Valle (2000); Baker e Kim (2004); Hambleton, Swaminathan e Rogers (1991); Klein (2003) e Pasquali (1997).



REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. da C. *Teoria da Resposta ao Item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística, 2000.

BAKER, F. B.; KIM, S. *Item response theory: parameter estimation techniques*. New York: Marcel Dekker, 2004.

BIRNBAUM, A. Some latent trait models and their models and their use in inferring an examinee's ability. In: LORD, F. M.; NOVICK, M. R. (Ed.). *Statistical theories of mental test scores*. [S.l.]: Addison-Wesley, 1968. p. 397-479.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *A redação no Enem 2020: cartilha do participante*. Brasília, DF: Inep, 2020.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Entenda sua nota no Enem: guia do participante*. Brasília, DF: Inep, [2012]. Disponível em: <https://download.inep.gov.br/educacao_basica/enem/guia_participante/2013/guia_do_participante_notas.pdf>. Acesso em: 2 set. 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *[Homepage do instituto]*. Brasília, DF, [2021]. Disponível em: <www.inep.gov.br>. Acesso em: 26 ago. 2021.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. *Fundamentals of item response theory*. California: Sage publications, 1991.

KLEIN, R. Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (Saeb). *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 11, n. 40, p. 283-296, 2003.

PASQUALI, L. *Psicometria: teoria e aplicações*. Brasília, DF: Editora da Universidade de Brasília, 1997.





(cc) BY-NC

VENDA PROIBIDA

