

EDUCAÇÃO & AVALIAÇÃO

UMA
HISTÓRIA DA
TRI:
ENTENDA
A RÉGUA DE
AVALIAÇÃO DO
ENEM

SUPERVISÃO PEDAGÓGICA RENATO JÚDICE
E CAMILA KARINO
DIREÇÃO DE ARTE ROB FRIEDE

UMA
PUBLICAÇÃO



geekie

SUMÁRIO

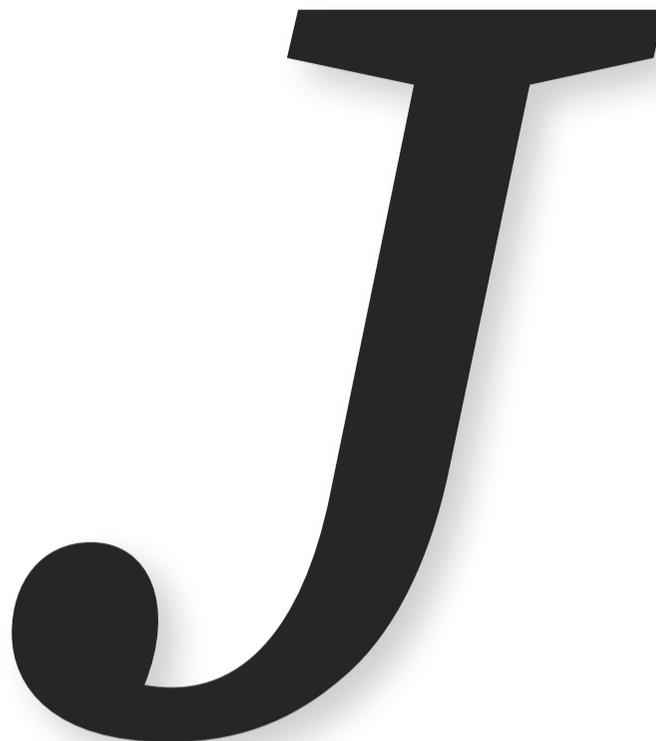
VOCÊ ENCONTRARÁ NESTE EBOOK:



INTRODUÇÃO	A TRI E A TRAJETÓRIA DE DUAS CANDIDATAS	3
PARTE 1	UMA COMPETIÇÃO DE SALTO EM ALTURA	5
PARTE 2	TODAS AS ETAPAS DO ENEM	9
PARTE 3	PRÉ-TESTE: QUAL É O PÚBLICO?	10
PARTE 4	QUEM TEM MEDO DA TEORIA ANTICHUTE?	11
PARTE 5	ZERO OU 300?	13
PARTE 6	USANDO A TRI NA SUA ESCOLA	15
PARTE 7	O FUTURO DO ENEM É DIGITAL	16

INTRODUÇÃO

A TRI É A TRAJETÓRIA DE DUAS CANDIDATAS



Julia é do tipo esguio. Tem 1,75 m e pesa 53 kg. Clarissa tem outro biótipo. Mede 1,68 m e pesa 64 kg. Julia usava um vestido leve azul-claro com estampas de flores e um par de sapatilhas e Clarissa vestia calça jeans com uma camiseta do Movimento Passe Livre e tênis quando as duas prestaram o primeiro dia de provas do Exame Nacional do Ensino Médio (Enem), em um sábado qualquer. Nossas duas personagens fictícias levam vidas bem diferentes e não se conhecem, mas têm o mesmo objetivo: conseguir vaga em um curso de Ciências Sociais em alguma universidade federal.

Filha de um psicólogo e de uma executiva, Julia meditou durante 10 minutos antes de sair de casa. Moradora de Pinheiros, zona oeste de

São Paulo, pegou carona com a mãe e seguiu de carro por 2,5 quilômetros até o local de prova. Filha de um pedreiro e de uma diarista, Clarissa rezou por 2 minutos no quarto de casa em Jandira, cidade próxima de Osasco, região oeste da Grande São Paulo. Depois pegou dois ônibus, em um trajeto de 8 km, para chegar à escola estadual onde prestou o Enem.

À 1 da tarde, Julia já estava sentada bem ereta na carteira da escola em Pinheiros, enquanto Clarissa ainda arrumava três copos de água embaixo da cadeira na sala em Jandira. Em minutos as duas abririam os cadernos de prova de Ciências Humanas e Ciências da Natureza e iniciariam uma “batalha” para mostrar o que aprenderam nos longos 12 anos de educação básica.

Os sonhos e o futuro de Julia e Clarissa dependem de algo que não pode ser observado ou medido diretamente por algum instrumento, ao contrário do peso e da altura de cada uma, do tempo que gastaram meditando ou rezando antes de saírem de casa ou da distância que percorreram para chegar ao local de prova. Eles dependem do conhecimento. Mas como medir o que uma pessoa sabe? Mais do que isso, como medir seu conhecimento comparado ao que ela deveria saber numa determinada época da vida escolar? E essa medida permite comparações mesmo sem utilizar o mesmo instrumento?

No caso do Enem, a alternativa correta para todas essas perguntas é a Teoria de Resposta ao Item (TRI), um sofisticado conjunto de modelos matemáticos. Julia e Clarissa não sabem, mas estão prestes a serem alocadas em uma régua, que vai dar a medida do seu domínio nas quatro áreas avaliadas no Enem: Ciências da Natureza, Ciências Humanas, Linguagens e Códigos e Matemática.

“A TRI serve para medir o traço latente, que é qualquer característica não só de pessoas, mas de empresas e outras instituições, que você só consegue medir de forma in-

direta”, diz Dalton Andrade, professor da Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina e um dos maiores especialistas em TRI do Brasil. Isso se aplica tanto ao que se aprende na escola como, por exemplo, ao nível de qualidade de vida de alguém. “Você monta um conjunto de perguntas e a partir das respostas consegue construir uma régua para medir aquela característica.” Resolver uma questão (item) sobre frações, por exemplo, exige o domínio específico de um conhecimento matemático.

A construção dessa régua só é possível porque Julia, Clarissa e os outros 8,5 milhões de candidatos que prestam o Enem com elas não são os pri-

meiros estudantes a responder às 180 questões que avaliam as quatro áreas do conhecimento. Meses ou até anos antes, grupos selecionados de alunos de escolas públicas e privadas de todo o Brasil já fizeram isso, em pré-testes realizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), órgão do Ministério da Educação que é o responsável pelo Enem, tanto do ponto de vista pedagógico quanto logístico.

Para os alunos do pré-teste, a pressão não era nem de perto a que Julia e Clarissa estavam sentindo naquele sábado, mas espera-se que eles tenham respondido àquelas questões com tanta seriedade quanto se valesse nota na escola. Para o Inep, era a

“ A TRI serve para medir o traço latente, que é qualquer característica não só de pessoas, mas de empresas e outras instituições, que você só consegue medir de forma indireta.”

Dalton Andrade, professor da Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina e especialista em TRI

oportunidade de descobrir o perfil das questões, para determinar, por exemplo, o lugar onde cada uma delas cairia na régua do exame. É o que se denomina tecnicamente de “calibrar o item”. Uma questão com grande porcentual de acertos no pré-teste é considerada fácil e fica localizada na parte inferior da régua. As questões de média dificuldade povoam o centro da régua e as mais difíceis são alocadas na parte superior da régua.

Continuando com nossas analogias com outros instrumentos de medida, posicionar as questões na régua de medida do conhecimento é tão importante quanto dar sentido às temperaturas auferidas na escala Celsius, por exemplo. Imagine que alguém menciona o valor de 40 graus centígrados. Como conhecemos bem essa escala, sabemos que esse valor equivale a um dia de sol muito, muito quente! Se alguém menciona 100 °C, sabemos que é a temperatura de ebulição da água ao nível do mar. E assim por diante...

Agora que o Inep posicionou os itens na régua do conhecimento, dado um determinado valor será possível reconhecer o que aquela pessoa é capaz de fazer (o que ela conhece), o que ela está aprendendo e o que ela ainda

não consegue fazer (o que ela ainda não conhece). Só que a dificuldade do item não é o único critério levado em conta, como veremos mais adiante.

Com o pré-teste, temos parte da régua pronta. A outra será dada pelos próprios candidatos. E aí surge logo a pergunta sobre nossas personagens. Onde Julia e Clarissa vão se encaixar na régua do exame? Será que mais para o início, denotando que elas sabem pouco? Ou será mais para o fim, sugerindo que elas são verdadeiras nerds, “sabem tudo”? Será que na régua de Matemática elas vão se posicionar bem, mas na de Ciências Humanas vão se dar mal? E mais, será que uma das duas terá uma posição mais privilegiada que a da outra? Enfim, as perguntas são inúmeras, mas todas poderão ser respondidas assim que tiver-

mos a posição delas nas régua, isto é, assim que soubermos a nota TRI de Julia e Clarissa no Enem. A essa nota dá-se o nome de proficiência.

PROVA DE SALTO EM ALTURA

Para entender bem o que é proficiência, vamos a outra analogia. Podemos imaginar que a trajetória das meninas lembre a de atletas de salto em altura. Julia e Clarissa saltaram sem problemas quando o “sarrafo” estava nas posições iniciais da régua de Ciências da Natureza. Mantiveram o ritmo, acertando as questões, até o sarrafo ficar na posição 500. Nesse ponto, Julia deixou de acertar. Clarissa, porém, seguiu adiante até o sarrafo ficar na posição 512,5. Veja abaixo uma amostra de questões e seus respectivos sarrafos, conforme mapa de itens divulgado pelo Inep:

JULIA



“SARRAFO” NA POSIÇÃO 465.

JULIA E CLARISSA SALTAM

Os carrinhos de brinquedo podem ser de vários tipos. Dentre eles, há os movidos a corda, em que uma mola em seu interior é comprimida quando a criança puxa o carrinho para trás. Ao ser solto, o carrinho entra em movimento enquanto a mola volta à sua forma inicial. O processo de conversão de energia que ocorre no carrinho descrito também é verificado em

- (A) um dínamo.
- (B) um freio de automóvel.
- (C) um motor a combustão.
- (D) uma usina hidroelétrica.
- (E) uma atiradeira (estilingue).

CLARISSA



“SARRAFO NA POSIÇÃO 485,9.

JULIA E CLARISSA SALTAM

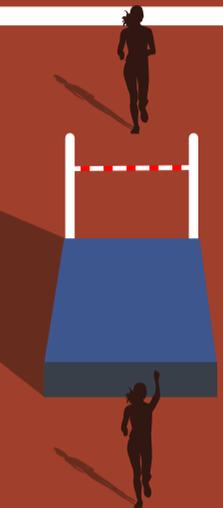
As estrelas-do-mar comem ostras, o que resulta em efeitos econômicos negativos para criadores e pescadores. Por isso, ao se depararem com esses predadores em suas dragas, costumavam pegar as estrelas-do-mar, parti-las ao meio e atirá-las de novo à água. Mas o resultado disso não era a eliminação das estrelas-do-mar, e sim o aumento do seu número.

DONAVEL, D. A bela é uma fera. Super Interessante. Disponível em: <http://super.abril.com.br>. Acesso em: 30 abr. 2010 (adaptado).

A partir do texto e do seu conhecimento a respeito desses organismos, a explicação para o aumento da população de estrelas-do-mar baseia-se no fato de elas possuírem

- (A) papilas respiratórias que facilitaram sua reprodução e respiração por mais tempo no ambiente.
- (B) pés ambulacrários que facilitaram a reprodução e a locomoção do equinodermo pelo ambiente aquático.
- (C) espinhos na superfície do corpo que facilitaram sua proteção e reprodução, contribuindo para a sua sobrevivência.
- (D) um sistema de canais que contribuíram na distribuição de água pelo seu corpo e ajudaram bastante em sua reprodução.
- (E) alta capacidade regenerativa e reprodutiva, sendo cada parte seccionada capaz de dar origem a um novo indivíduo.

JULIA



“SARRAFO” NA POSIÇÃO 500.
JULIA PASSA E CLARISSA FICA

Pesticidas são contaminantes ambientais altamente tóxicos aos seres vivos e, geralmente, com grande persistência ambiental. A busca por novas formas de eliminação dos pesticidas tem aumentado nos últimos anos, uma vez que as técnicas atuais são economicamente dispendiosas e paliativas. A biorremediação de pesticidas utilizando microrganismos tem se mostrado uma técnica muito promissora para essa finalidade, por apresentar vantagens econômicas e ambientais. Para ser utilizado nesta técnica promissora, um microrganismo deve ser capaz de

- (A) transferir o contaminante do solo para a água.
- (B) absorver o contaminante sem alterá-lo quimicamente.
- (C) apresentar alta taxa de mutação ao longo das gerações.
- (D) estimular o sistema imunológico do homem contra o contaminante.
- ~~(E)~~ metabolizar o contaminante, liberando subprodutos menos tóxicos ou atóxicos.

~~CLARISSA~~



“SARRAFO” NA POSIÇÃO 502,3.
JULIA SALTA



Suponha que, em cena anterior à apresentada, o homem tenha se alimentado de frutas e grãos que conseguiu coletar. Na hipótese de, nas próximas cenas, o tigre ser bem-sucedido e, posteriormente, servir de alimento aos abutres, tigre e abutres ocuparão, respectivamente, os níveis tróficos de

- (A) produtor e consumidor primário.
- (B) consumidor primário e consumidor secundário.
- ~~(C)~~ consumidor secundário e consumidor terciário.
- (D) consumidor terciário e produtor.
- (E) consumidor secundário e consumidor primário.

JULIA



“SARRAFO” NA POSIÇÃO 512,5. JULIA FICA



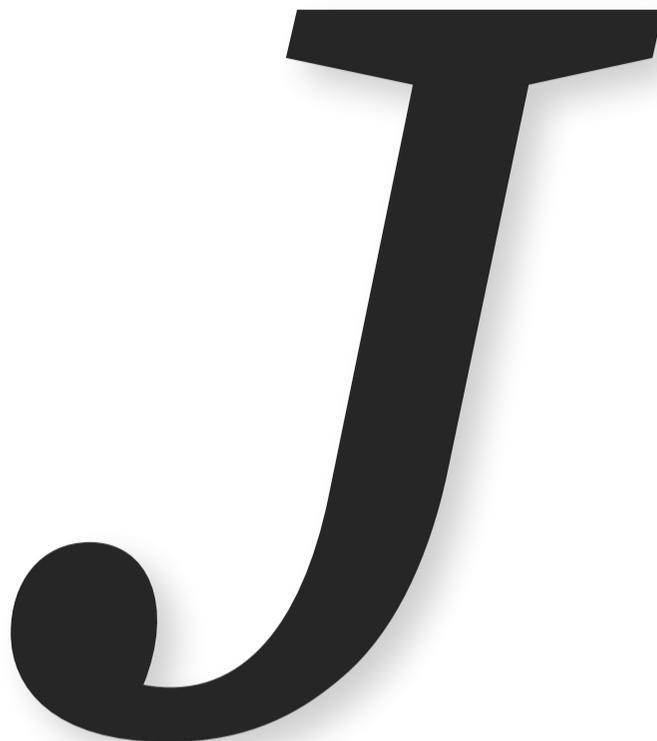
CLARISSA

De acordo com o relatório “A grande sombra da pecuária” (“Livestock’s Long Shadow”), feito pela Organização das Nações Unidas para a Agricultura e a Alimentação, o gado é responsável por cerca de 18% do aquecimento global, uma contribuição maior que a do setor de transportes.

Disponível em: www.conpet.gov.br. Acesso em: 22 jun. 2010.

A criação de gado em larga escala contribui para o aquecimento global por meio da emissão de

- (A) metano durante o processo de digestão.
- (B) óxido nítrico durante o processo de ruminação.
- (C) clorofluorcarbono durante o transporte de carne.
- (D) óxido nítrico durante o processo respiratório.
- (E) dióxido de enxofre durante o consumo de pastagens.



Julia termina o primeiro dia de provas e avisa à mãe pelo WhatsApp que está liberada. Enquanto espera, Clarissa, que terminou de preencher o gabarito cinco minutos depois, começa a caminhar em direção ao ponto de ônibus. Vai conversando com uma vizinha que encontrou no local de prova, que será sua companhia no trajeto de volta para casa.

TODAS AS ETAPAS DO ENEM

Enquanto nossas atletas de salto em altura seguem para um merecido descanso, na expectativa do segundo dia de exames, temos uma boa oportunidade para falar, em ordem cronológica, de todas as fases do processo de preparação do Enem. A primeira delas é justamente a elaboração de questões como essas que as meninas acabaram de responder, uma das etapas mais compli-

cadas da maratona de montar uma prova com base na TRI.

Avaliações baseadas em questões e não no conjunto da prova, como a TRI, exigem um grande volume de itens de qualidade, capazes de aferir o desempenho de candidatos. Com uma rede extensa de professores como a envolvida no Enem, em tese não seria difícil pinçar os 180 itens usados em cada exame. Não é bem assim. “Quando se começou a trabalhar com TRI no Brasil, nos anos 90, de cada seis questões você aproveitava uma”, diz Dalton.

A situação melhorou, mas ainda é alto o descarte de itens inadequados. Muitas vezes eles não servem porque não distinguem bem quem sabe de quem não sabe; noutras porque não tratam do que o aluno deve dominar (competência) para passar no Enem. Esses domínios, como “aplicar as tecnologias da

comunicação e da informação na escola, no trabalho e em outros contextos relevantes para sua vida”, estão descritos na matriz de referência do exame.

Para analisar esse último aspecto, a triagem das questões tem uma etapa pedagógica, que ocorre antes do pré-teste. Nessa etapa, muitas questões são descartadas e não são aproveitadas na elaboração dos instrumentos de medida. A quantidade de questões descartadas depende de características das avaliações (estilo de questão, matriz e etapa de ensino) e da experiência dos professores-elaboradores.

A construção de um banco de itens que dê suporte às avaliações é um desafio. Mesmo exames mais antigos, como o Sistema de Avaliação da Educação Básica (Saeb), que já usa a TRI desde 1995, tiveram dificuldades com a consolidação de um banco de questões, diz Amaury Gremaud, professor de Economia da Universidade de São Paulo em Ribeirão Preto e diretor de Avaliação do Inep de 2006 a 2008.

O Enem passou a usar a TRI bem depois do Saeb, em 2009. Desde então, o Inep acelerou o processo de elaboração de questões, fechando uma parceria com 37 universidades e institutos federais de ensino. Mesmo assim, ninguém de fora do governo se arriscar a estimar o tamanho do Banco Nacional de Itens.

PRÉ-TESTE: QUAL É O PÚBLICO?

Já falamos dessa etapa antes, mas vale a pena retomar o tema, e não só para manter a ordem

PASSO A PASSO NA ELABORAÇÃO DE ITENS TRI

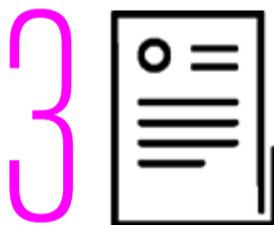


1 Baseado nas características de cada questão, o Enem exige a elaboração de um grande volume delas. Uma rede de professores prepara os itens, muitos deles descartados já na primeira triagem pedagógica.



2

No pré-teste, grupos restritos de alunos respondem a lotes de questões que serão incluídas em edições futuras do Enem. As com grande índice de acerto vão para a base da régua do exame, as de média dificuldade povoam o centro e as difíceis, o topo.



3 Na análise estatística, as questões são avaliadas pelo grau de discriminação, de dificuldade e de chance de acerto ao acaso. Cada variável dessas gera uma curva. Quanto antes ela sobe, por exemplo, mais fácil é a questão.

4

Questões reprovadas nas avaliações são descartadas de vez ou encaminhadas para reelaboração.



5



Questões aprovadas na triagem vão para o Banco Nacional de Itens, que abastece exames oficiais.

6

A cada ano, são pinçadas 180 questões do banco para compor o Enem. As notas mínima e máxima dependem dos parâmetros de dificuldade dos itens, por isso não há nota 0 nem nota 1000.



cronológica dos passos necessários à preparação do Enem. Como vimos, é com base nas respostas de grupos selecionados de estudantes que o Inep monta a régua que mais tarde vai medir o desempenho real dos candidatos no Enem. Mas, assim como a fase de elaboração das questões, o pré-teste pode ser problemático.

Um dos aspectos em jogo é o público do pré-teste. É preciso ter uma amostra equilibrada de estudantes do ensino médio de todo o Brasil. “Tem um problema com os pré-testes, que é o fato de muitas vezes eles serem feitos por alunos sem tanta proficiência. Acontece muito com matemática. Você fica com um monte de itens que em tese não prestam, mas o problema está em quem responde”, diz Dalton. “A solução seria pegar também alunos de escolas públicas e privadas de bom desempenho.”

Especialistas admitem ainda que há uma diferença de performance entre o público descompromissado do pré-teste e os candidatos que estão jogando seu futuro no Enem. “Só que a desmotivação no pré-teste não afeta o comportamento do item. Uma questão de dificuldade 600 vai manter a posição na escala. O que pode ocorrer é a proficiência do aluno ser subavaliada”, diz Tufi Machado Soares, pro-

“ Em geral, se o processo de elaboração das questões está bem afinado, o descarte de itens no pré-teste será baixo, na faixa de 5% a 10%.”

Camila Karino, diretora de Avaliação da Geekie e ex-coordenadora-geral de Instrumentos e Medidas do Inep

fessor do Departamento de Estatística e do Programa de Doutorado e Mestrado em Educação da Universidade Federal de Juiz de Fora.

“O problema de o aluno responder de qualquer forma porque o pré-teste não conta na nota existe. Mas no Saeb, por exemplo, os erros são sempre iguais, é sistemático”, diz o matemático Ruben Klein, consultor da Cesgranrio, presidente da Associação Brasileira de Avaliação Educacional e pioneiro na aplicação da TRI no Brasil com o Saeb de 1995 – na época, ele fez até o programa de computador que rodou a prova. “Você vê um padrão de respostas se repetindo e quando posiciona os itens numa escala aquilo faz um sentido enorme.”

Ao mesmo tempo que precisa montar uma amostra abrangente, o pré-teste tem de ser limitado ao mínimo possível de escolas, por questões de segurança.

QUEM TEM MEDO DA TEORIA ANTICHUTE?

Chegar ao pré-teste não garante que uma questão vá ser incorporada ao Banco Nacional de Itens. Após serem respondidos em uma aplicação do pré-teste, os itens são analisados estatisticamente. “Em geral, se o processo de elaboração das questões está bem afinado, o descarte de itens no pré-teste será baixo, na faixa de 5% a 10%”, diz Camila Karino, diretora de Avaliação da Geekie e ex-coordenadora-geral de Instrumentos e Medidas do Inep.

Nessa fase de análise, busca-se verificar empiricamente a qualidade da questão. São realizadas análises clássicas e estimam-se os parâmetros dos itens, conforme o modelo da TRI utilizado no Enem. Uma das características analisadas é o poder de discriminação, ou seja, a capacidade do item de separar quem domina o conteúdo de quem não o domina. “Digamos que no Enem caia uma questão simples, de soma, e 99,9% dos alunos acertem. Essa questão não traz nenhuma informação para a avaliação”, diz Tufi. “Por outro lado, se aplicada no 5º ano do ensino fundamental, ela pode ser útil.”

A segunda característica avaliada é a dificuldade. Na régua do Enem, a posição de uma questão está relacionada ao ní-

vel de proficiência exigido dos candidatos para que eles assinalem a resposta correta.

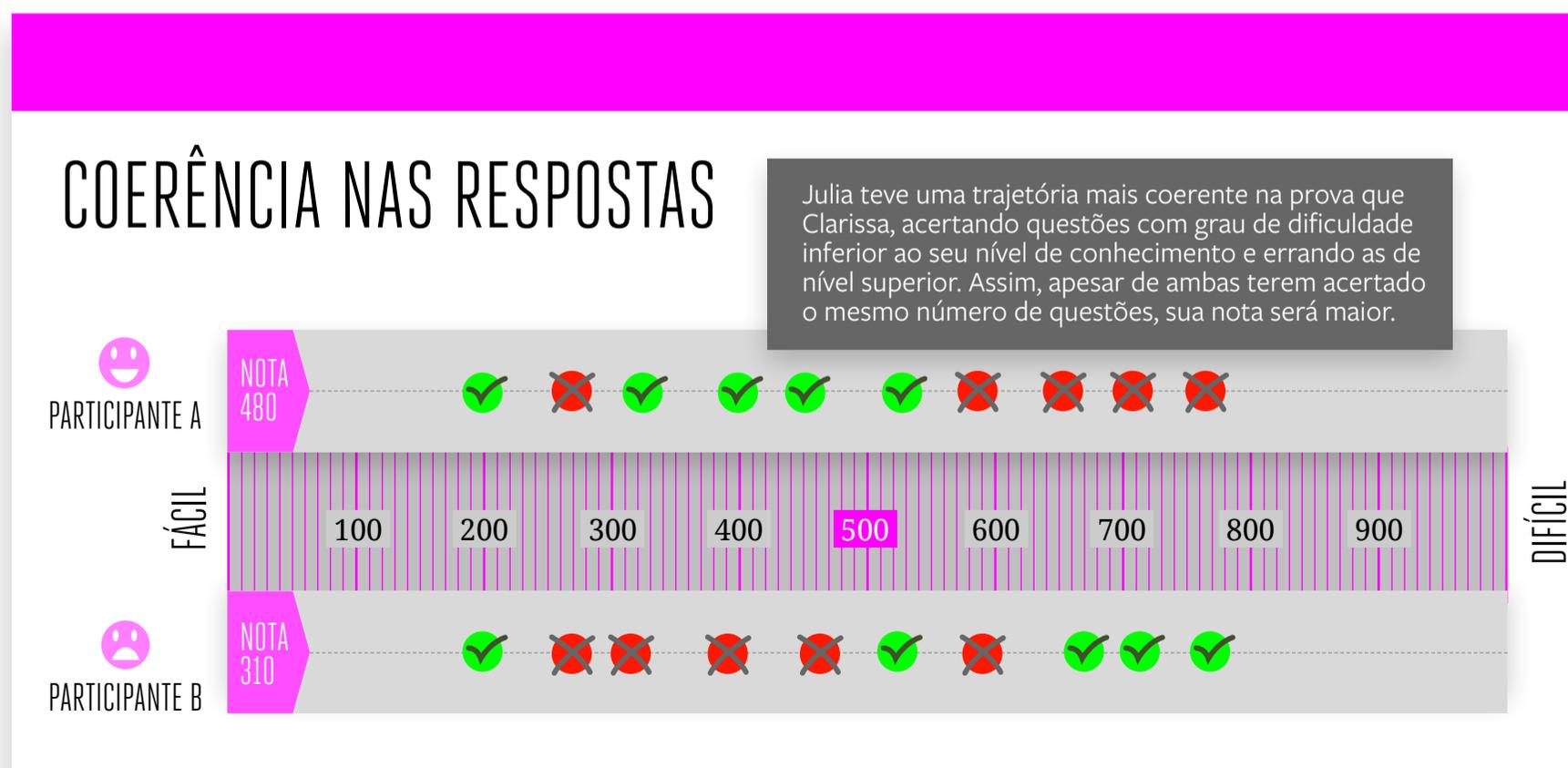
O terceiro parâmetro é o que busca estimar se um determinado aluno acertou a resposta na base da adivinhação, do chute. Essa variável já valeu ao Enem o apelido de “Teoria Antichute” e confundiu estudantes, como veremos um pouco mais adiante. Ela tenta estimar a capacidade de alguém assinalar a resposta correta mesmo que não saiba o necessário sobre um determinado assunto. Afinal, o candidato tem diante de si cinco alternativas e pelo menos 20% de chances de acertar a resposta ao acaso.

“A expressão chute está errada conceitualmente, porque todos têm uma probabilidade mínima de acertar”, afirma Tufi,

que prefere usar a expressão inglesa *guessing* (algo como palpite) para designar esse tipo de comportamento dos candidatos. Uma hipótese que aumenta a chance de *guessing* bem-sucedido é a de as alternativas erradas (distratores) serem muito óbvias. “A pessoa pode acertar por um processo simples de eliminação.”

Durante a análise de questões, os técnicos do Inep observam as curvas características dos itens (CCI) como a apresentada abaixo. A régua do Enem tem como referência 500, que foi a nota média no Enem 2009 dos alunos que estavam concluindo o ensino médio naquele ano.

Pelas curvas é possível analisar todas as características das questões. “Quanto antes



a curva subir, mais fácil é o item”, explica Camila. Ou seja, usando como referência o gráfico abaixo, quanto mais a curva se desloca para a direita, mais difícil é a questão. Por outro lado, curvas mais eretas indicam maior poder de discriminação entre alunos que sabem e não sabem. “Se ela fica mais deitada, esparramada, você sabe que a questão tem menor poder de separar o joio do trigo”, diz Amaury.

NOTAS E O ZERO QUE VIRA 300

O cálculo das notas do Enem é outra operação complexa, rea-

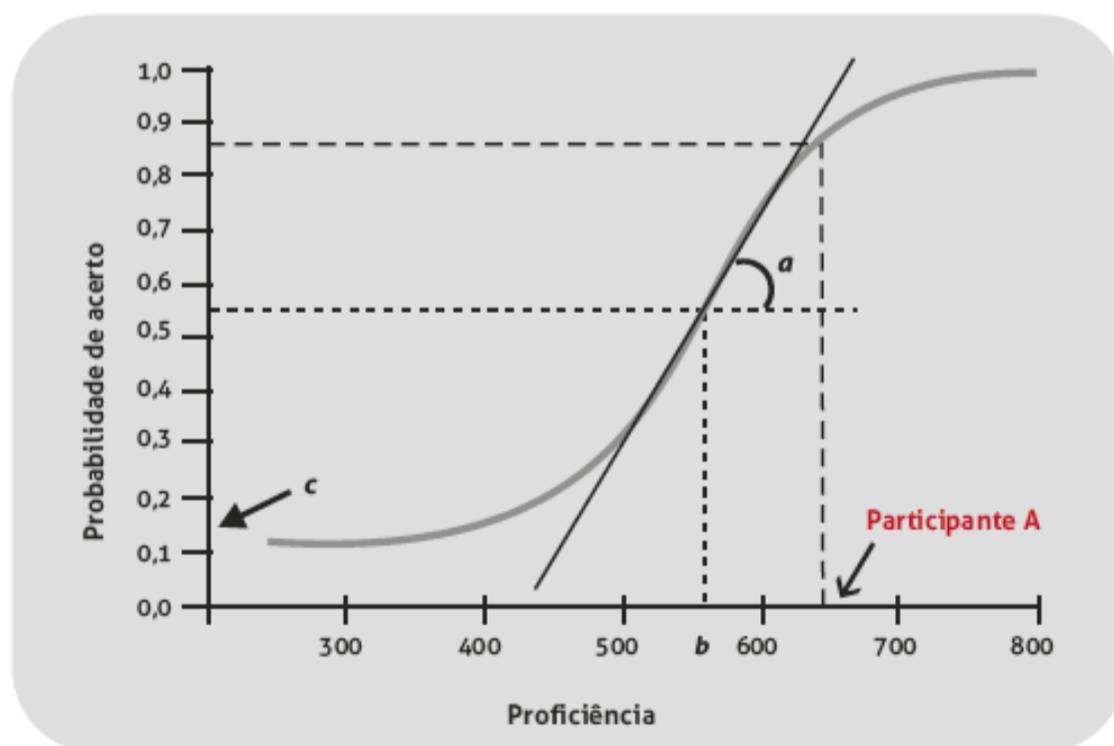
lizada por equipes que reúnem especialistas em estatística, matemática ou psicometria, área que mistura as duas primeiras com a psicologia.

O processo é todo complicado, mas o que candidatos como Julia e Clarissa – por sinal, as duas acabaram de iniciar a prova de Linguagens e Códigos no segundo dia do exame – e suas famílias querem mesmo entender são dois pontos: por que ninguém recebe nota zero no Enem? E por que estudantes que acertaram o mesmo número de questões podem ter pontuações diferentes? Vamos a eles:

A nota mínima e máxima do Enem estão relacionadas aos parâmetros de dificuldades dos itens que compõem a prova. No exame de 2014, em Matemática, por exemplo, a nota mínima foi de 318,5. Ou, em Linguagens e Códigos, de 306,2. Esses eram os menores valores de proficiência que podiam ser estimados dadas as características dos itens daquele ano. Da mesma forma, a nota máxima em Matemática foi de 973,6 – a de Linguagens e Códigos ficou um pouco abaixo, em 814,2 – ou seja, se o candidato acertasse todas as questões da prova, esses seriam os valores de proficiência (nota) que eles receberiam.

O professor Dalton, da Federal de Santa Catarina, brinca com a preocupação com o zero. “Essa cultura de que quem não sabe nada tira zero, coitado do zero, ele tem seu valor. Só que no Enem o zero às vezes vira 300.” Ou seja, no Enem o aluno que erra todas as questões da prova tira a nota mínima, que não necessariamente é o zero.

Já a diferença de notas entre quem acertou o mesmo número de questões obedece a outro padrão, o da coerência. Considerando que o conhecimento é cumulativo e o domínio de habilidades mais complexas requer o domínio de habilidades mais simples, espera-se que o estudante acerte



Os técnicos do Inep analisam gráficos como este para saber quanto uma questão consegue discriminar quem sabe de quem não sabe um conteúdo (a), seu grau de dificuldade (b) e a possibilidade de acerto na base do chute (c). Quanto antes a curva subir, por exemplo, mais fácil é o item; uma curva mais ereta indica maior capacidade de discriminação. “Se ela fica mais esparramada, você sabe que a questão tem menor poder de separar o joio do trigo”, diz Amaury Grenaud, ex-diretor de Avaliação do Inep

as questões com nível de dificuldade abaixo do seu nível de conhecimento e tenda a errar as questões que exigem um nível de conhecimento superior ao que ele possui. Assim, acertos aleatórios, não coerentes, tendem a valer menos na estimação da proficiência do candidato.

No caso das nossas candidatas, Julia teve um padrão de acertos mais coerente pedagogicamente do que Clarissa. Assim, apesar de as duas terem acertado o mesmo número de questões, a nota de Julia será superior.

Aqui é preciso fazer um parêntese. A divulgação da “Teoria Antichute” pela imprensa atribuiu a essa variável um peso que ela não tem. Não se trata de uma metodologia punitiva que desconta pontos se a questão é acertada ao acaso. Por isso, não se justifica a atitude de outros candidatos, que preferiram deixar em branco questões que não sabiam com medo de perder nota. Acertos sempre contam a favor.

Enquanto Julia abre o caderno de prova de Matemática e Clarissa resolve a questão número 44 de Linguagens e Códigos, você pode estar se perguntando: por que o Inep adota um processo tão trabalhoso no Enem? Por que não levar em conta apenas o número de acer-

tos, como é feito em exames nas escolas e até mesmo em concursos públicos ou grandes vestibulares, como a Fuvest?

Esse modelo, do número de acertos, é o da Teoria Clássica dos Testes. Essa é a teoria mais comumente utilizada, mas ela tem limitações que foram superadas pela TRI. Para começar, a nota na Teoria Clássica depende do grau de dificuldade da prova. “Você pode ter um aluno que tira 7 em um exame. Mas em outro exame, se você aumentar o grau de dificuldade, ele tira 4. E aí, esse aluno é nota 4 ou nota 7?”, pergunta o especialista em avaliação Renato Júdice.

Outra característica da TRI, tão importante quanto essa, é a de que provas diferentes são comparáveis. O exame da Teoria Clássica tem como foco a prova. E a prova como um todo nunca vai ser igual a outra (paralelas). A TRI, ao ter como base uma escala de referência, possibilita a construção de provas compa-

ráveis mesmo que elas não sejam 100% paralelas.

É importante ressaltar, porém, que estamos falando em comparabilidade, não em provas de dificuldade idêntica. Isso porque é praticamente impossível ter dois lotes de 45 questões com o mesmo grau de dificuldade. O próprio Inep admitiu que o Enem de 2011 “tinha questões relativamente mais fáceis” do que o exame de 2010. Só que, como em ambos os casos os itens estavam posicionados na mesma régua, o único efeito prático foi que os candidatos de 2011 tiveram de acertar mais questões para conseguir a mesma pontuação dos que prestaram o exame anterior.

Ressalvas à parte, a comparabilidade é importante por dois motivos. O primeiro é garantir justiça na avaliação, mesmo que com provas diferentes. Poucos se dão conta, mas o Enem já é realizado duas vezes por ano, primeiro para os candidatos tra-

dicionais e depois para presos e internos de unidades socioeducativas, como a Fundação Casa.

Também fazem a prova nessa segunda data candidatos que, por um motivo ou por outro, ficaram sem nota no primeiro exame. Em diversas edições do Enem isso ocorreu com moradores de áreas que ficaram inundadas no dia da primeira prova. Outro grupo que fez o Enem na segunda data foi o de 639 estudantes do Ceará que teve as provas anuladas pelo vazamento do pré-teste em 2011.

O segundo motivo pelo qual a comparabilidade é importante vem do fato de o Enem não ser só um processo seletivo. Um dos objetivos do exame é fazer um raio X da educação básica ao término do ensino médio, fornecendo dados para orientar políticas públicas. Isso não seria possível se o exame não fosse comparável ano a ano, para sa-

ber se as políticas adotadas têm dado algum resultado.

USANDO A TRI NA SUA ESCOLA

Graças em boa parte à TRI, o Enem é um instrumento confiável de mensuração do desempenho dos estudantes brasileiros. O problema é que a montanha de dados gerados raramente tem servido para aperfeiçoar o ensino na prática. E não estamos falando só no nível do macro, dos governos estaduais e federal, mas também no das escolas. O Enem poderia ser um roteiro do que é necessário para formar estudantes qualificados, mas essa chance tem sido, na maioria dos casos, pouco aproveitada.

Especialista em avaliação da Geekie, Claudio Maroja, colaborador do Banco de Itens do Inep, fala com essa experiência e com a de professor das redes municipal e estadual de São

Paulo. “É preciso usar o resultado dos exames de forma a fazer efeito. No Brasil você tem muitos dados, mas eles não são de fácil leitura”, diz. “O Inep vê a avaliação como fim, não como meio. Para nós (educadores) ela é meio.”

Traduzir o resultado da avaliação em ação pedagógica é uma tarefa complexa. É preciso montar um plano de ensino e uma sequência didática, saber o que o aluno domina e o que é necessário ele aprender, etapa por etapa, para chegar a uma pontuação maior. Não se trata de apagar incêndios, mas criar um ciclo de trabalho.

Muitas vezes, mesmo nas escolas privadas, que em tese têm melhor estrutura, o uso dos resultados das avaliações deixa a desejar. Uma reação comum aos resultados do Enem é adotar a política de apertar o ritmo de estudos para fazer a nota subir no exame seguinte. Isso ocorre, por exemplo, quando a escola cai no ranking do exame e passa a temer a perda de alunos para concorrentes.

“Como puxar a nota é mais fácil com alunos de maior proficiência, as escolas esquecem quem está em um nível inferior”, diz Maroja. “Chegam ao ponto de fazer provas específicas só para turmas avançadas ou só para os piores alunos. Elas se esquecem de que o estudante

“É preciso usar o resultado dos exames de forma a fazer efeito. No Brasil você tem muitos dados, mas eles não são de fácil leitura. O Inep vê a avaliação como fim, não como meio. Para nós (educadores) ela é meio.”

Claudio Maroja, especialista em avaliação da Geekie e colaborador do Banco Nacional de Itens

não é de nível inferior, ele está em um nível inferior.”

Uma das saídas para essa situação pode ser a escola incorporar a TRI e as provas de resultados comparáveis ao seu dia a dia. A Geekie fornece uma plataforma que utiliza a TRI, numa escala comparável à do Enem, para calibrar questões preparadas por uma rede própria de professores. Como a nota é comparável à do exame oficial, as avaliações permitem antecipar o desempenho dos estudantes e, com isso, possibilitar o delineamento da melhor estratégia de ensino. De modo a contribuir com o processo de ensino, a plataforma da Geekie também aponta os pontos fracos e fortes do aluno e monta para ele um plano de estudos, atualizado periodicamente graças a algoritmos de ensino adaptativo.

Dissemos lá no começo do texto que a TRI é um conjunto de modelos matemáticos. Alguns deles são mais complexos que o do Enem e permitem a realização de avaliações com outros propósitos e características.

Por exemplo, há modelos da TRI que permitem a utilização do chamado item politômico ordinal, ou seja, admite respostas graduais do tipo discorda/discorda totalmente/concorda/concorda totalmente.

Também é possível montar provas que consideram acer-

tos parciais. Conhecido como Modelo de Samejima em homenagem à sua criadora, a psicometrista japonesa Fumiko Samejima, essa variante da TRI já foi usada no Brasil, na análise da prova de escrita da ANA (Avaliação Nacional de Alfabetização).

ENEM ONLINE, A APOSTA PARA O FUTURO

No médio prazo, entretanto, a mudança mais provável do Enem é sua transformação em prova digital. Em janeiro o ex-ministro da Educação, Cid Gomes, acenou com a possibilidade de adotar um exame online. Os candidatos marcariam a data na qual fariam o exame, como já ocorre com testes internacionais. É o caso do Toefl, exame que avalia conhecimentos em inglês, tem várias edições por ano e é realizado em mais de 4.500 locais de 165 países.

O próprio Gomes reconheceu na época que, antes de mais nada, é preciso ampliar bastante

o número de questões do Enem armazenadas no Banco Nacional de Itens. O crescimento do banco também é requisito para o desenvolvimento de provas adaptativas. Nelas, o computador submete as questões ao aluno de acordo com a resposta ao item anterior. O objetivo é ter um vestibular o mais personalizado possível.

Amaury Gremaud, da USP Ribeirão, concorda que o Enem online traria vários benefícios, começando por eliminar o pesadelo logístico que é organizar o exame hoje. “Se você coloca 8 milhões de pessoas sentadas para fazer uma prova em um só dia país afora, há um grande risco de que haja problemas”, diz. “Só que para ter o Enem online você precisa de um banco de itens fenomenal. Os grandes exames lá de fora têm centenas de milhares de questões.”

Para Ruben Klein, da Cesgranrio, o exame online é mesmo o futuro, desde que o

banco de itens seja realmente robusto. “Nos Estados Unidos tiveram de tirar do ar um teste adaptativo porque as questões estavam muito repetitivas”, diz Klein. “Aqui no Brasil tem cursos de preparação para testes do Detran em que você vai lá treinar e as mesmas questões caem na prova.”

Klein acredita, no entanto, que para o Enem digital dar certo são necessárias mudanças na lei, com o objetivo de resguardar o sigilo das questões. “O Enem é um tipo de concurso no qual o aluno sai com a prova. O ideal numa prova adaptativa é que não tivesse uma legislação que te obrigasse a divulgar o exame”, diz.

Dalton Andrade, da Federal de Santa Catarina, diz que, apesar das ressalvas, a prova adaptativa é a ideal, porque permite medir o conhecimento do candidato de forma mais precisa sem que ele precise enfrentar uma maratona de questões. “Na prova única você tem pessoas de níveis muito diferentes, quem está lá em baixo tem muito pouco item para acertar e quem está lá em cima tem muito pouco item para errar”, diz. “É um tipo de exame no qual você tem problemas nessas duas pontas.”

Na prova adaptativa, o candidato pode responder a um maior número de questões

que ficam no seu nível de conhecimento ou próximas dele. “O aluno só vai responder aos itens difíceis se for de alta proficiência, caso contrário vai ficar nas questões mais fáceis. É preciso ter um número significativo de itens para que a pessoa possa se expressar bem neles”, explica Dalton.

Parece cronometrado. Mal o professor Dalton acabou de falar e Julia terminou a redação. Deu um suspiro de alívio e alisou o vestido – que desta vez tinha motivos geométricos no lugar de flores. Dali a alguns meses, ela conferiu pela internet sua pontuação no Enem. Foi bem o suficiente para entrar no curso de Ciências Sociais. Feito bastante comemorado em casa, com os pais, e mais tarde com amigos num barzinho da Vila Madalena.

Clarissa terminou a redação quase meia hora depois de Julia.

Teve mais dificuldades nas provas objetivas do exame, mas foi muito bem na redação. No final, ela também conseguiu nota suficiente para entrar na educação superior e realizou o sonho do pai.

O fato de Clarissa ter conseguido entrar na faculdade graças à redação mostra a importância dessa prova na composição da nota final. Essa é a única prova do Enem que não usa a Teoria de Resposta ao Item. O Inep já convocou universidades e institutos de pesquisa interessados em investigar meios de adotar a TRI na produção de texto. “Estão sendo feitos estudos. Uma possibilidade é a de o corretor entrar como mais uma métrica de avaliação, de se tentar determinar o quanto um corretor é mais rigoroso do que outro”, diz Camila, da Geekie. Mas a redação já é assunto para um próximo ebook.

CADA PESSOA APRENDE DE UM JEITO DIFERENTE. ENTÃO, POR QUE ENSINAR DA MESMA FORMA?

A Geekie é uma solução educacional capaz de potencializar o aprendizado e melhorar o desempenho de cada aluno sem deixar ninguém para trás



COM A GEEKIE, SEUS ALUNOS E A SUA ESCOLA PODEM MUITO MAIS

RELATÓRIOS GERENCIAIS

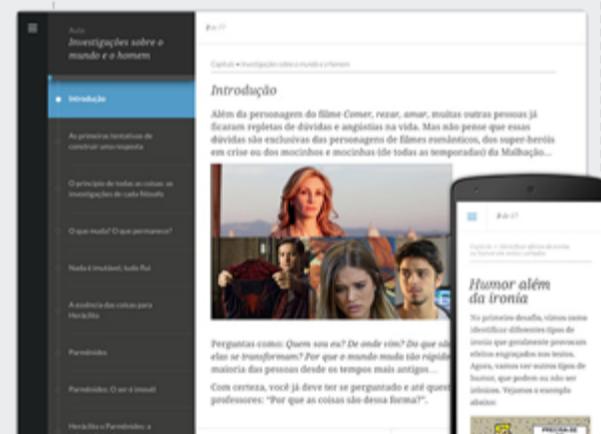
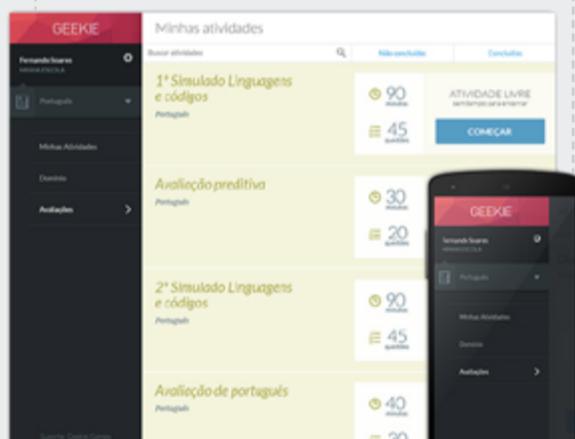
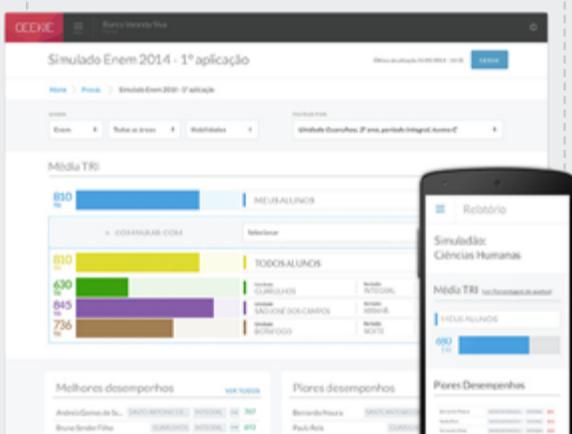
Professores e gestores têm acesso a relatórios de desempenho de cada aluno e, assim, conseguem fazer intervenções direcionadas

PREPARAÇÃO PARA O ENEM

Avaliações na escala ENEM no começo e fim do ano letivo. O aluno tem uma estimativa da sua nota no exame oficial e tem um poderoso instrumento de análise da evolução

ENGAJAMENTO DO ALUNO

Conteúdo dinâmico, desenvolvido para a plataforma digital



INTERESSADO?

A gente vai adorar falar com você

SAIBA MAIS

www.geekie.com.br