



iede

Interdisciplinaridade e Evidências no Debate Educacional



lede Pedagógico

Sumário

Parte I - Distribuições populacionais

1. Curva Normal, Desvio-Padrão, escala única e distorções na distribuição

Parte II - Teoria de Avaliação

1. Entendendo o que é Média e a sua aplicação na Teoria Clássica dos Testes (TCT)
2. Entendendo o que é Curva Característica do Item e sua aplicação na Teoria de Resposta ao Item (TRI)
3. Histogramas

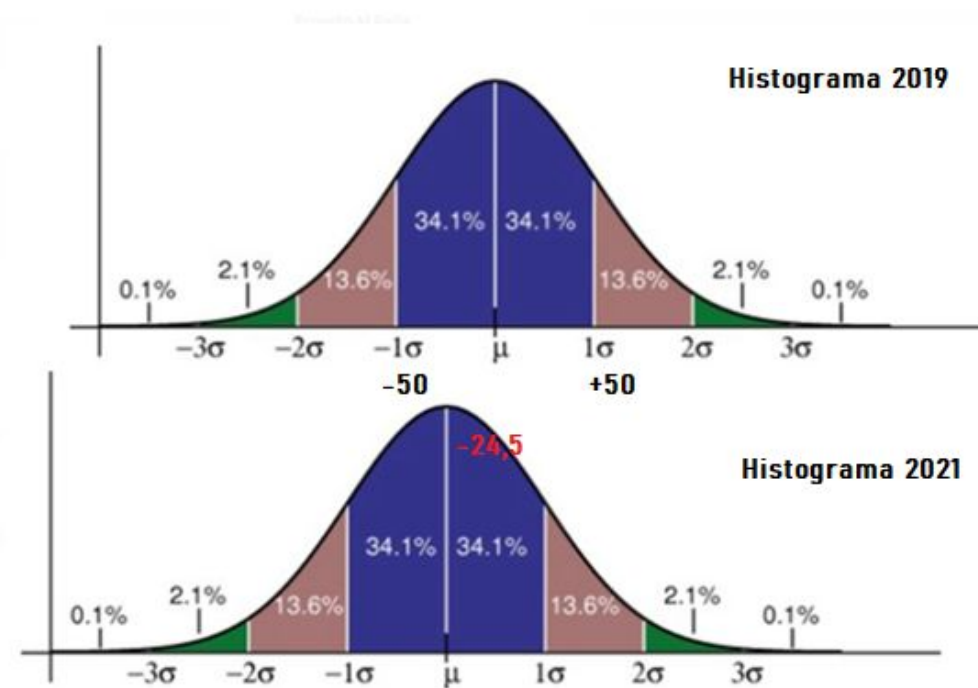
Parte III - Discussão sobre o que é um bom item e apresentação do Iede Pedagógico

1. Revisando Exemplos de itens bons x itens ruins
2. Entendendo o que são distratores e o que significam os erros dos estudantes

Distribuições populacionais

Curva Normal

- Curva Normal: conceito estatístico que descreve graficamente a distribuição de dados em torno de uma média, onde a maioria dos valores se concentra próximo à média, e se distanciam gradualmente à medida que se afastam deste ponto central. Imagine que você tem todas as notas de um teste e as coloca em um gráfico. Aqui está como funciona:
- A **Média**: No centro do gráfico, você tem uma linha que representa a média das notas. A média é basicamente a nota "típica" que a maioria das pessoas conseguiu.
- A **Concentração**: A maioria das notas está agrupada perto dessa média. Isso significa que a maioria das pessoas teve um desempenho parecido.
- As **Diferenças**: À medida que você se afasta da média, as notas começam a se espalhar. Algumas pessoas vão ter notas mais altas e se destacar, enquanto outras vão ter notas mais baixas e ter mais dificuldades.

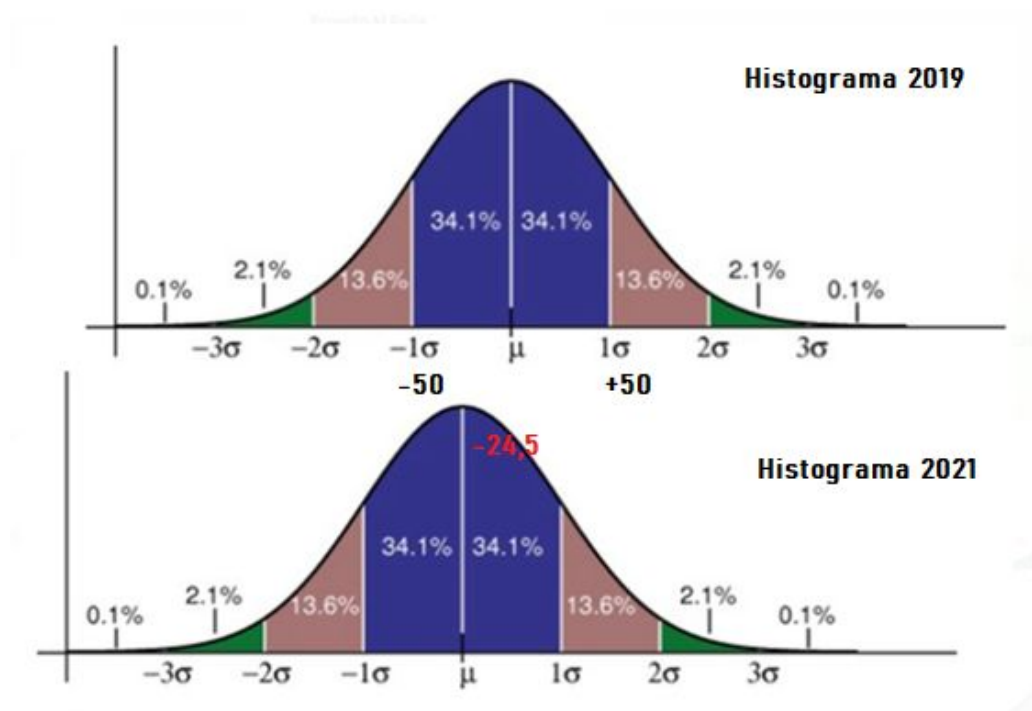


Fonte: Saeb 2019 e Saeb 2021.

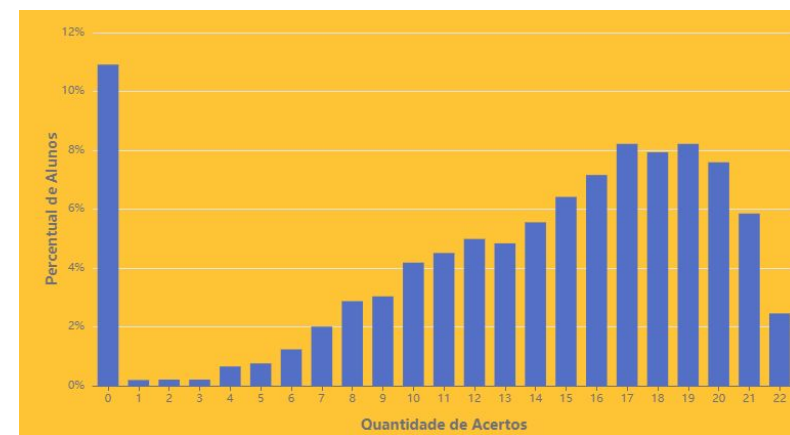
Curva Normal

Quando olhamos para a distribuição das notas, podemos tirar algumas conclusões:

- Se a distribuição das notas **se parece muito com essa curva normal**, isso geralmente sugere que o **teste conseguiu discriminar bem os alunos de alto desempenho e os alunos de baixo desempenho**.
- Se há uma concentração de resultados em **valores acima de média** pode significar que **faltaram itens mais difíceis na avaliação**.
- Se há uma concentração de resultados em **valores abaixo da média** pode significar que **faltaram itens mais fáceis na avaliação** ou que há um grupo de estudantes com um nível muito abaixo do esperado.



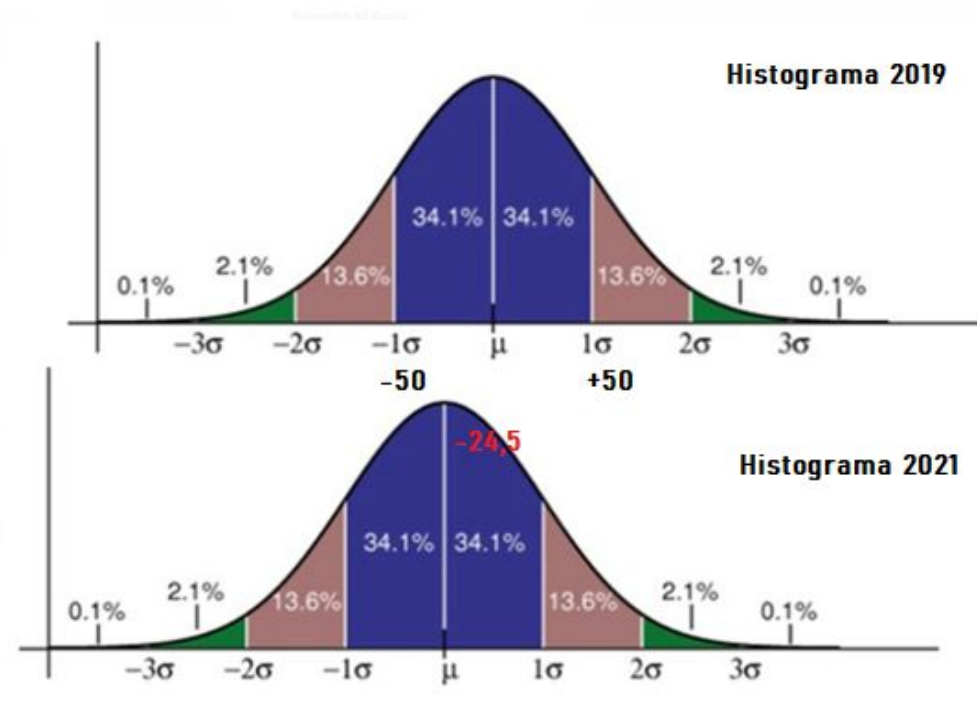
Fonte: Saeb 2019 e Saeb 2021.



Desvio-padrão

Desvio-Padrão: é uma ferramenta que nos ajuda a ver o quão espalhadas ou próximas as notas dos alunos estão da nota média do teste.

- O Ponto Central é a nota média que a maioria dos alunos alcançou em um teste. O Desvio-Padrão ajuda a medir o quão longe cada nota individual está dessa média. Ele nos diz a distância média entre as notas de todos os alunos e a nota média.
- Quanto **maior o Desvio-Padrão**, mais distantes estão as notas dos alunos da média, ou seja, **há uma grande diferença entre as notas mais altas e as mais baixas**.
- Se o **Desvio-Padrão for pequeno**, isso significa que as notas dos alunos estão mais próximas da média, e **há menos diferença entre as notas mais altas e mais baixas**.



Fonte: Saeb 2019 e Saeb 2021.

Sobre Escala única para diferentes idades e/ou séries

Tabela de peso e altura por idade, para meninos de 3 a 18 anos de idade

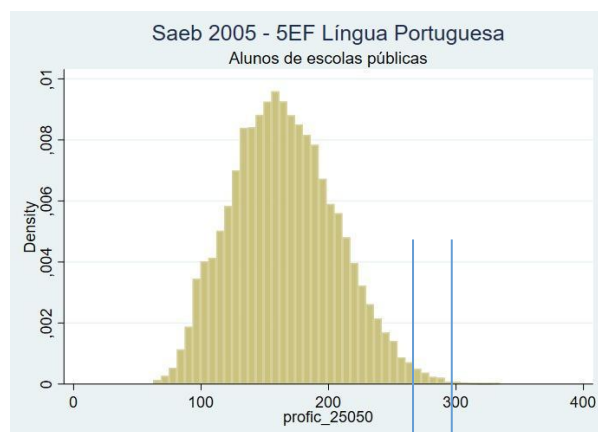
Idade (anos)	Mínimo		Médio		Máximo	
	Peso (kg)	Altura (cm)	Peso (kg)	Altura (cm)	Peso (kg)	Altura (cm)
3	12,25	90,6	14,61	96,2	17,78	102,8
4	13,65	97,5	16,51	103,4	20,09	110,4
5	15,24	102	18,37	108,7	22,86	117,1
6	17,46	108,5	21,91	117,5	27,71	126,2
7	19,50	114	24,54	124,1	31,71	133,4
8	21,77	119,6	27,26	130	36,02	140,2
9	23,81	124,2	29,94	135,5	40,73	145,3
10	25,76	128,7	32,61	140,3	45,36	150,3
11	28,03	133,4	35,20	144,2	50,67	154,4
12	30,48	138,1	38,28	149,6	56,34	161,5
13	32,66	142,2	42,18	155	62,60	169,5
14	36,20	146,4	48,81	162,7	68,31	177,1
15	41,41	151,7	54,48	167,8	73,30	181,8
16	46,90	156,5	58,83	171,6	77,34	185,6
17	50,33	160	61,78	175,7	80,05	189,5
18	53,33	163,5	64,33	179,5	82,33	193,5

Tabela de peso e altura por idade, para meninas de 3 a 18 anos de idade

Idade (anos)	Mínimo		Médio		Máximo	
	Peso (kg)	Altura (cm)	Peso (kg)	Altura (cm)	Peso (kg)	Altura (cm)
3	11,61	88,4	14,42	95,7	18,96	103,5
4	13,25	95,2	16,42	103,2	21,86	112,3
5	14,56	100	18,37	109,1	23,95	118,8
6	16,87	108	21,09	115,9	26,63	125,4
7	18,73	114	23,68	122,3	30,63	131,7
8	20,55	119,1	26,35	128	35,79	137,4
9	22,27	123,6	28,94	132,9	40,78	143,4
10	24,13	127,7	31,89	138,6	46,22	149,3
11	26,26	132,3	35,74	144,7	51,21	157,4
12	28,85	137,8	39,74	151,9	57,92	164,5
13	32,75	143,7	44,95	157,1	64,55	168,5
14	37,69	148,2	49,17	159,6	68,40	170,7
15	40,37	150,2	51,48	161,1	70,40	171,6
16	41,64	150,8	53,07	162,2	71,53	172
17	42,59	151	54,02	162,5	72,35	172,2
18	42,87	151	54,39	162,5	72,89	172,2

Pontos de atenção na análise e interpretação dos resultados do Ideb

- 5º ano do EF é o mais difícil de analisar só com as médias do Ideb

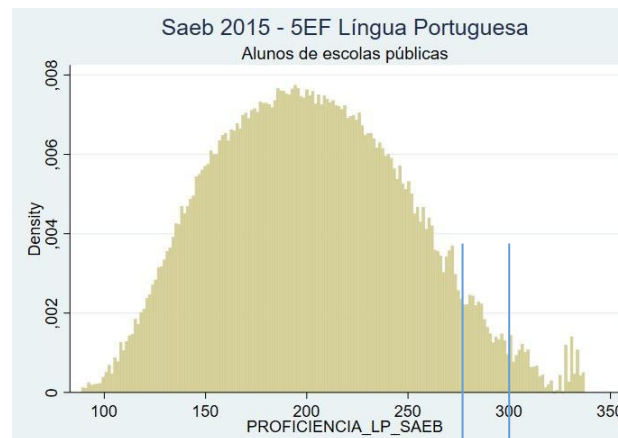


Pontuações consideradas como de aprendizado adequado na escala Saeb:

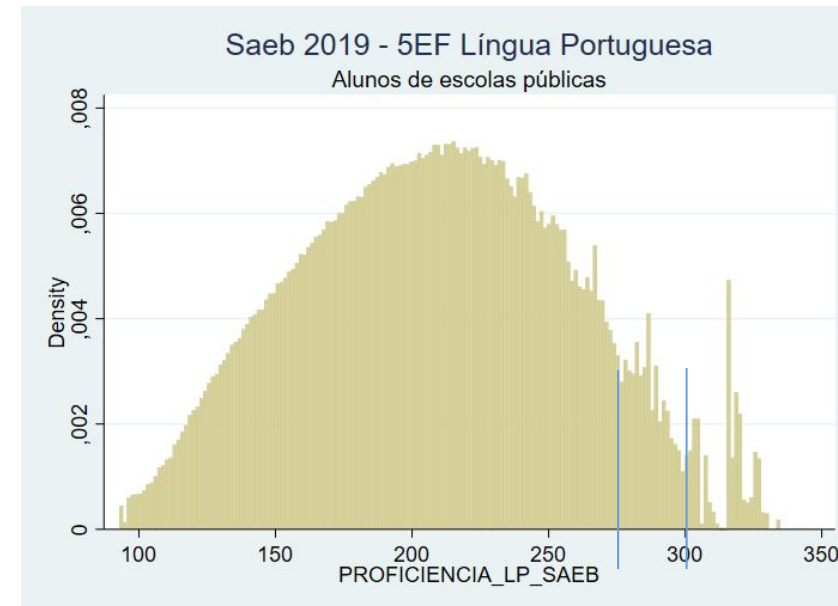
5º ano: 200

9º ano: 275

3º ano do EM: 300



Em 2015 (e mais ainda em 2019), há um número importante de alunos do 5º ano com aprendizado em LP considerado adequado ao 9º ano e até mesmo ao 3º ano do EM, o que é improvável na prática.



Distribuição estranha, com muitos alunos com resultados extremos

Mesma coisa que ter mais gente de 2m de altura do que de 1,90m numa população.

Teorias de Avaliação

TCT x TRI: principais diferenças

PRINCIPAIS CARACTERÍSTICAS		
Comparações	TCT	TRI
Nome da medida	Percentual de acerto, ou soma de pontos simples.	Proficiência - Medida de desempenho.
O que avalia	Avalia o processo de aprendizagem e pode proporcionar uma medida desses processos 'in loco'.	Produz uma medida de desempenho que leva em consideração toda a trajetória escolar do aluno.
Unidade de Interesse	O teste.	Os itens.
Comparabilidade	Limitada; para preservar a comparabilidade, deve ser utilizado o mesmo teste.	É possível variar os testes, desde que existam 'itens em comum'.

Limitações da Teoria Clássica dos Testes (TCT)

- Desempenho medido com base na **média** das avaliações;
- Incapacidade de medir competências individuais;
- Mais suscetível ao acaso (i.e. “chutes”);
- Subjetividade na elaboração estrutural do teste (conexão item-habilidade);
- Subjetividade na definição da **dificuldade** dos itens do teste;
- Pressupostos de **estabilidade** e **unidimensionalidade** da habilidade medida.

Média e a sua aplicação na Teoria Clássica dos Testes (TCT)

- Média: medida estatística que descreve a tendência central de um conjunto de dados. Na TCT, a média é usada para estimar o valor médio de uma característica ou habilidade medida por um teste.

$$\text{Média} = \frac{\text{Soma de todos os dados}}{\text{Tamanho da amostra}}$$






- A média é combinada com outras medidas, como o desvio padrão, para fornecer uma compreensão mais completa da distribuição dos escores.

Exemplo 1 - Média de idades:

Vamos considerar a idade de 5 alunos de uma turma. Para encontrar a média de idade da turma, somamos as idades dos alunos e dividimos pelo número de alunos:

$$\frac{(6 + 8 + 9 + 8 + 9)}{5} = \frac{40}{5} = 8 \text{ anos}$$

Portanto, a média de idade dos alunos é de **8 anos**.






Aluno	Idade
	6
	8
	9
	8
	9

Exemplo 2 - Média de notas:

Suponha que 6 alunos tiraram as seguintes notas em uma prova de Matemática: 8, 9, 6, 7 e 10. Para encontrar a média das notas, somamos todas elas e dividimos pelo número de alunos:

$$\frac{(8 + 9 + 6 + 7 + 10)}{5} = \frac{40}{5} = 8$$

Portanto, a média das **notas dos alunos** é **8**.

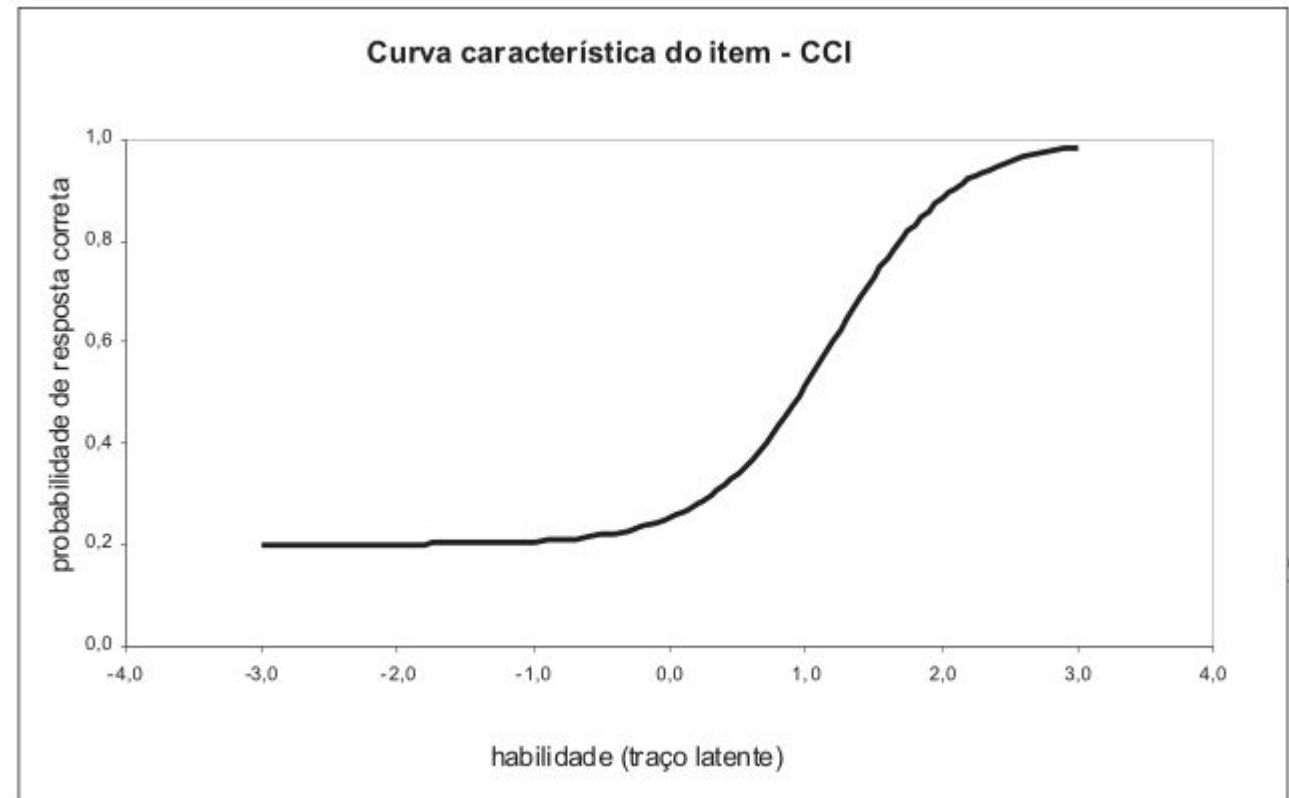
Aluno	Nota
	8
	9
	6
	7
	10

Teoria de Resposta ao Item (TRI)

Nome	1. Você, quando está no avião ou no ônibus, consegue colocar a mala no compartimento de cima?	2. Seu corpo não cabe dentro de um colchão tamanho padrão?	3. Você conversa com a maioria das pessoas olhando ao menos ligeiramente para baixo?	4. Você acredita que, pela sua altura, poderia se dar bem em jogos de basquete e vôlei contra seus amigos	5. Você tem altura para enterrar uma cesta de basquete?	6. Você frequentemente ajuda pessoas a colocar ou retirar objetos de armários altos?
Lais Costa	1	0	1	1	0	1
Leandro Mauri	1	1	1	1	0	0
Marta	1	1	0	0	0	1
Diva	1	0	0	0	0	0
Fatima	1	0	0	0	0	0
MARY	1	0	0	0	0	0
ADALBERTO	1	0	0	1	1	1
Marilu	1	0	0	1	0	0
José Aparecido	1	0	0	0	0	0
Bruno	1	0	0	0	0	1

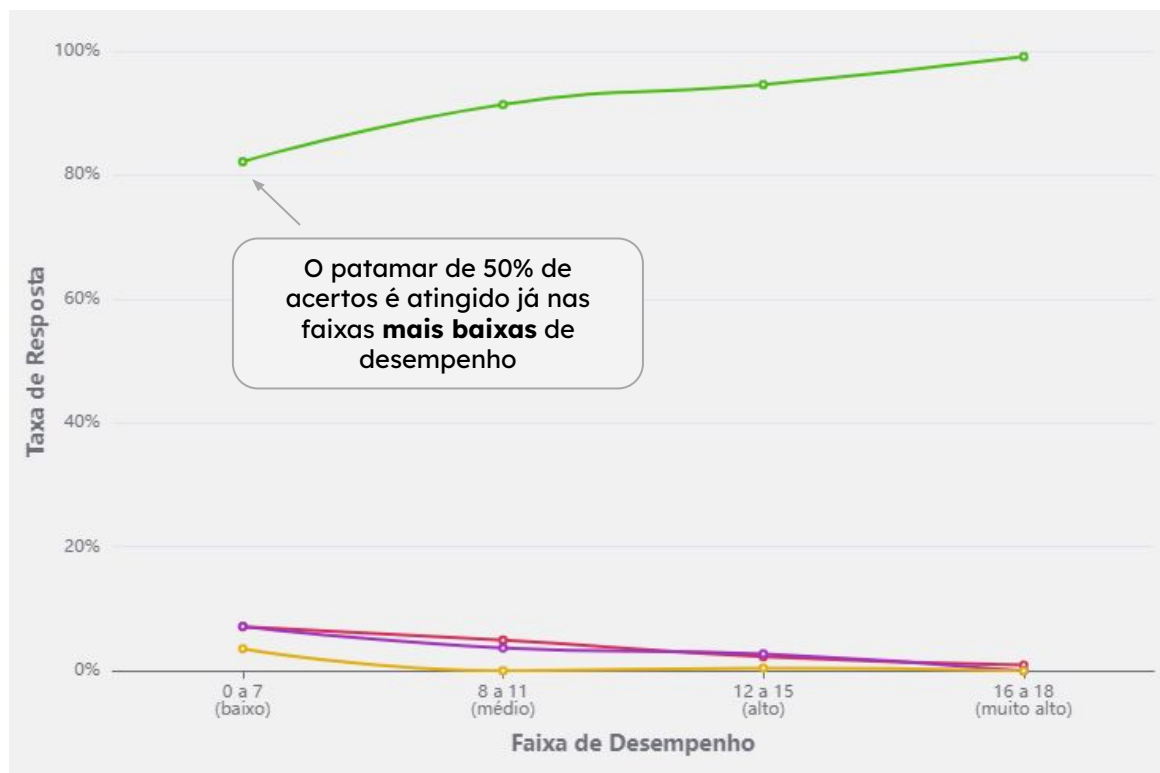
Curva Característica do Item (CCI) e sua aplicação na Teoria de Resposta ao Item (TRI)

- A CCI é uma representação gráfica da probabilidade de acerto, que varia em função da **habilidade** do indivíduo em cada item.
- Com ela, podemos quantificar a **dificuldade** de um item, que pode ser representada pelo ponto na CCI em que a probabilidade de acerto é de 50%, ou seja, onde a curva atinge seu ponto médio.
- Já a **discriminação** de um item é indicada pela inclinação da CCI. Quanto mais inclinada a curva, maior é a capacidade do item de diferenciar entre indivíduos com diferentes níveis de habilidade.

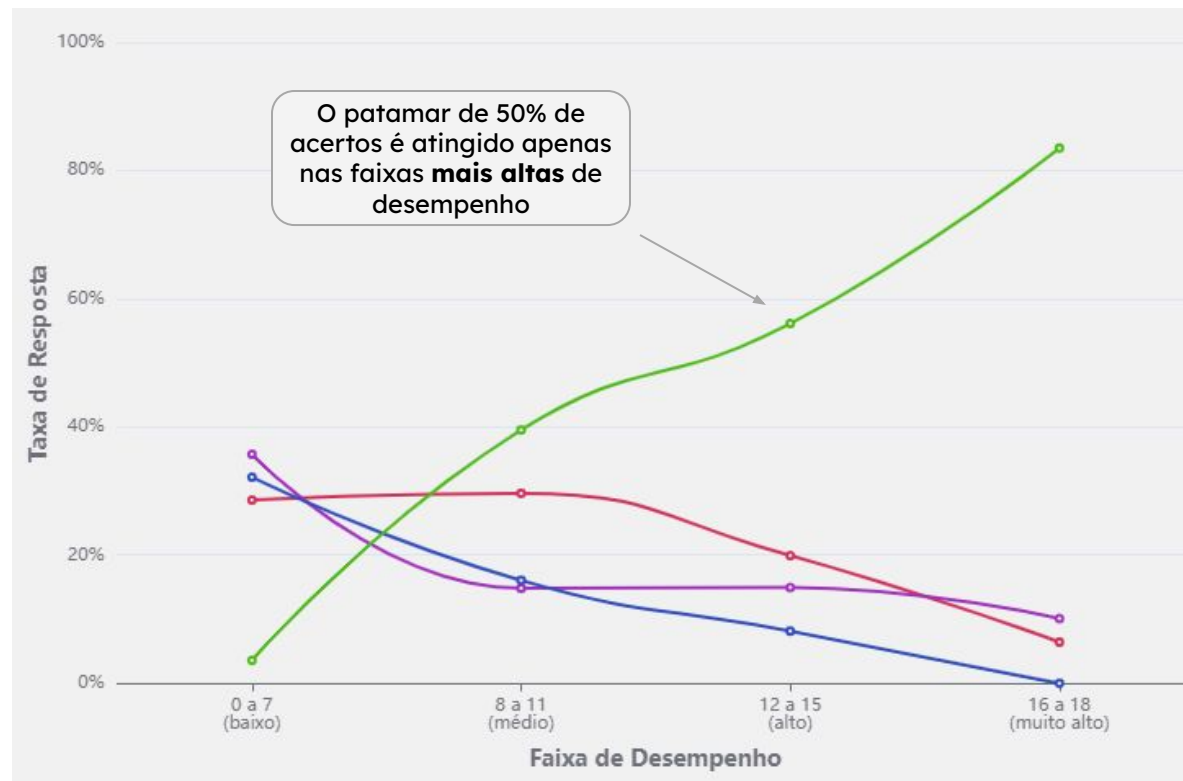


CCI como parâmetro de dificuldade

CCI - Questão **fácil**

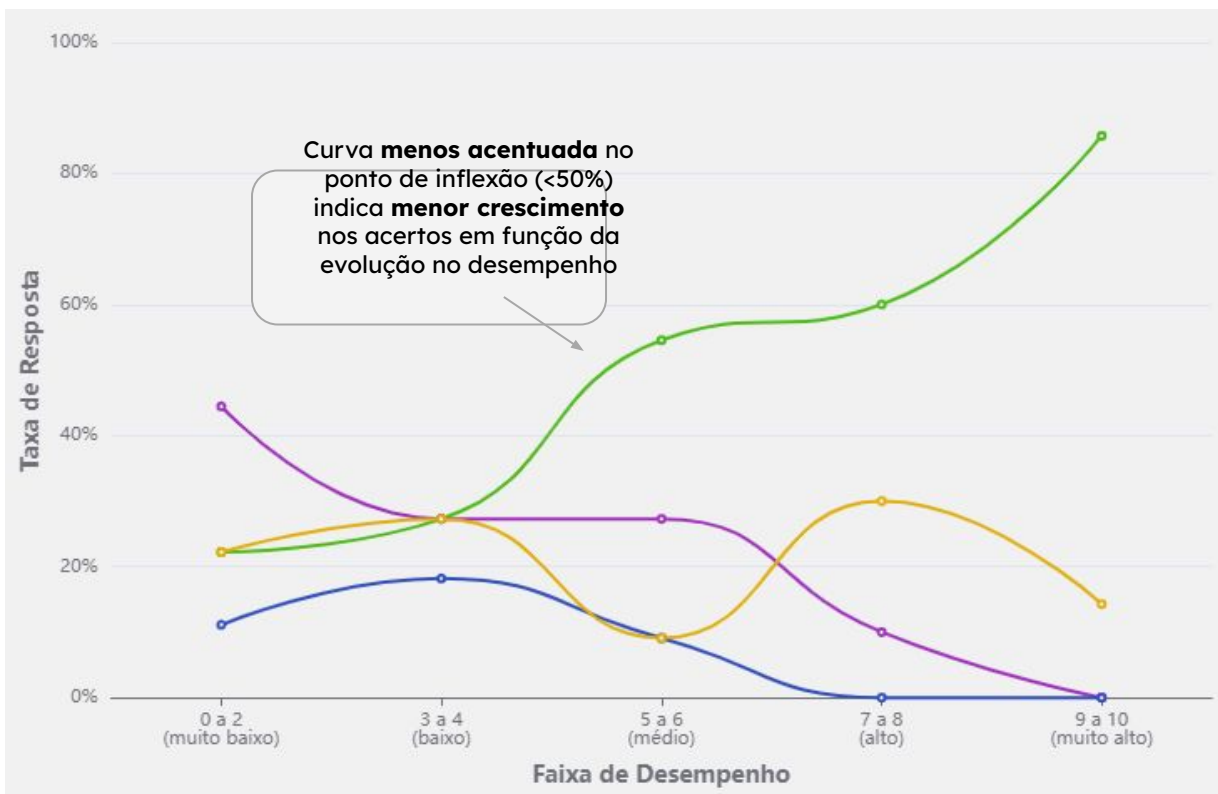


CCI - Questão **muito difícil**

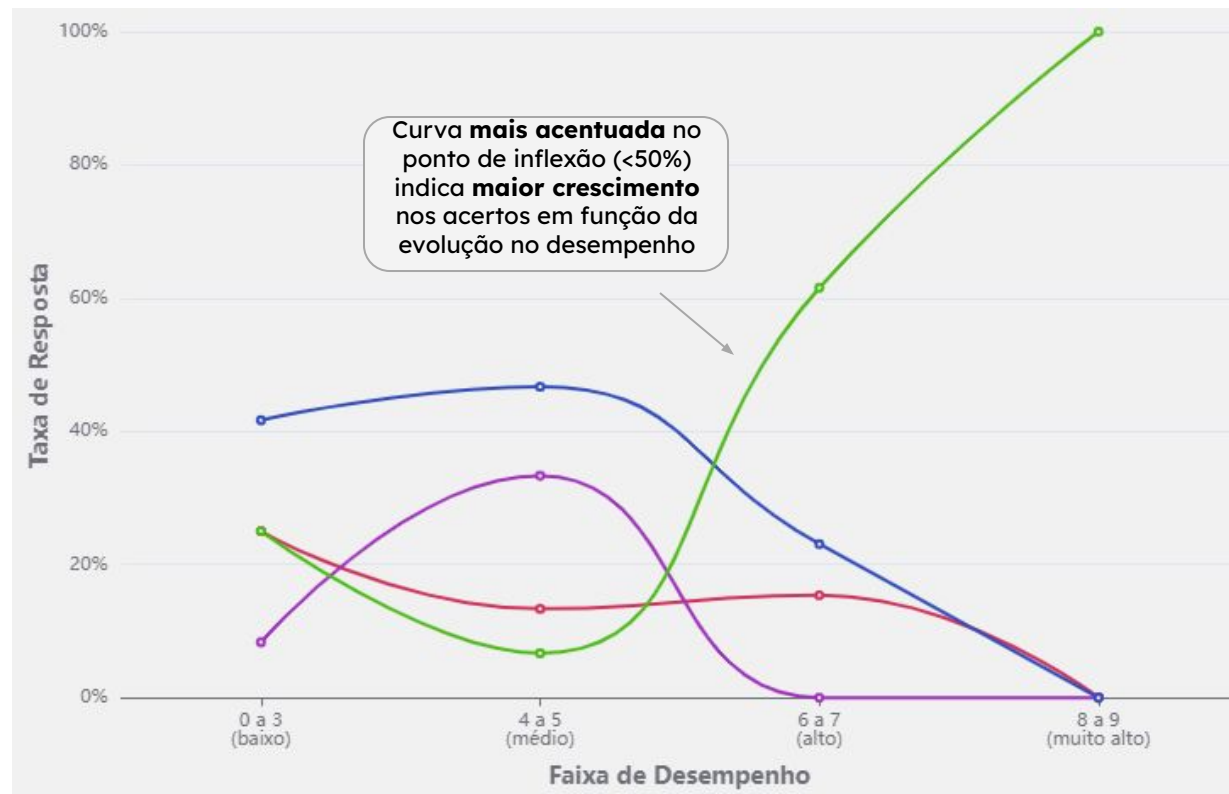


CCI como parâmetro de discriminação

CCI - Questão **pouco** discriminativa

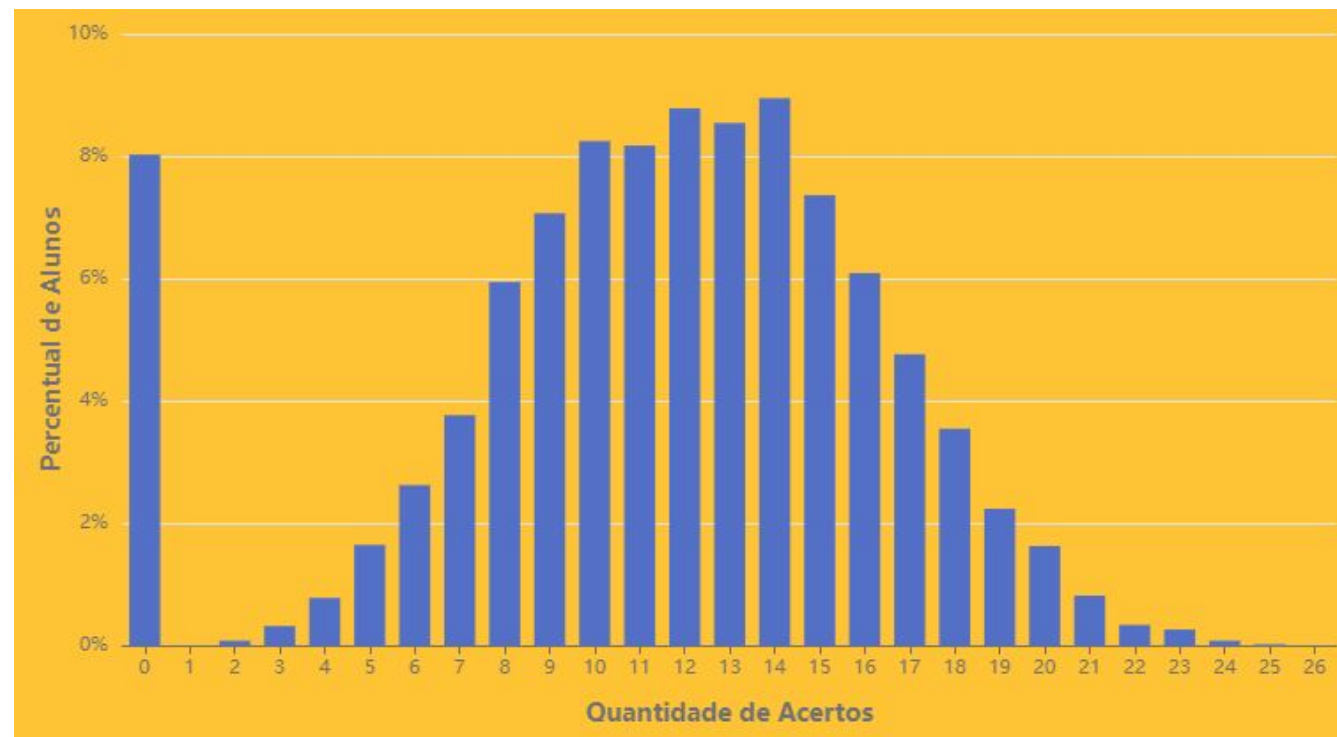


CCI - Questão **muito** discriminativa



Histogramas

- Histograma: **representação gráfica da distribuição dos dados**
- É usado para visualizar a **frequência** com que os **valores** ocorrem em um **conjunto de dados**.
- Ajuda a identificar padrões: assimetrias, presença de picos, etc.
- Também auxilia na detecção de valores atípicos (outliers) e na compreensão da concentração ou dispersão dos dados.



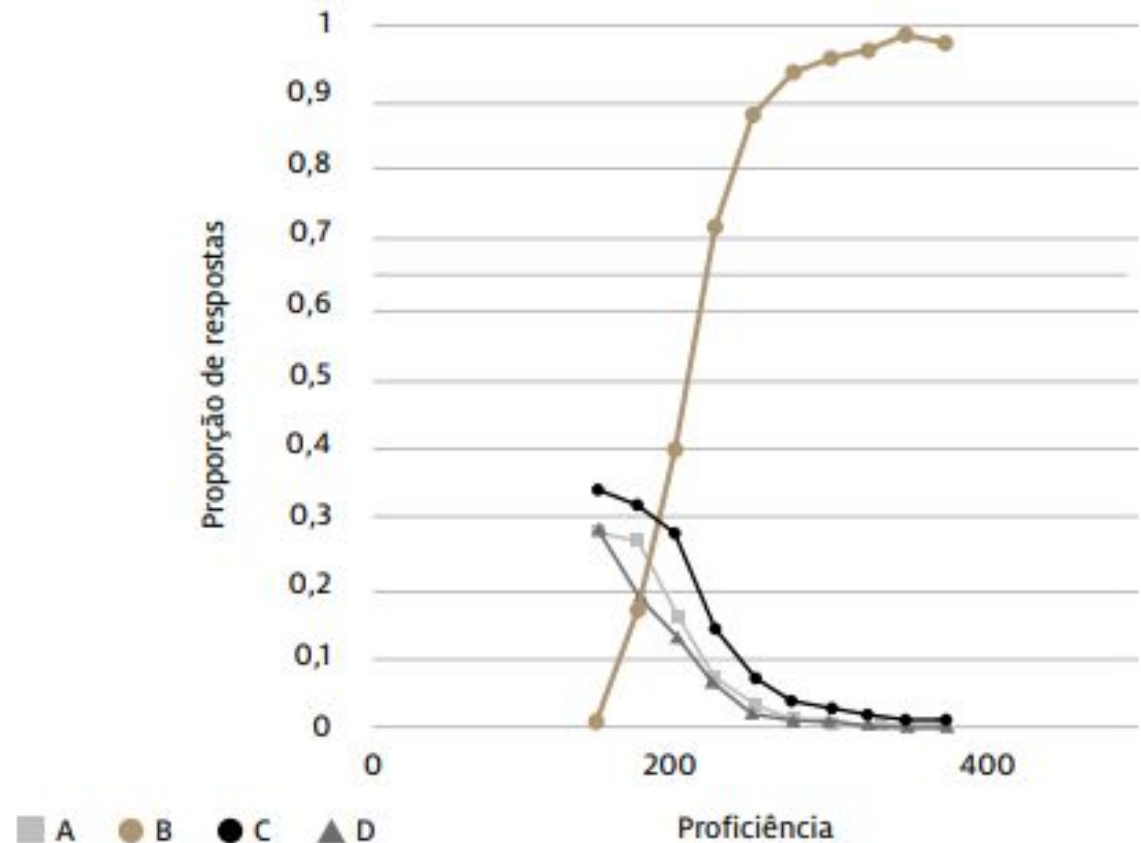


O que é um bom item?

Item bom x item ruim

- **Item bom:** item que, individualmente, aponta **resultados distintos** para **estudantes com diferentes níveis de domínio de uma competência**.
- O Gráfico ao lado ilustra o percentual de acertos dos estudantes em um item aplicado no Saeb, em que o eixo horizontal indica o nível de proficiência dos alunos, e o eixo vertical, a proporção de respostas dadas em cada alternativa, sendo que a correta é a B (representada pela linha laranja).
- Esse gráfico revela como se comporta um **bom item, que consegue discriminar adequadamente os estudantes:** quanto mais elevada sua proficiência, mais alto o índice de escolha da resposta certa

PROPORÇÃO DE ESCOLHA DAS ALTERNATIVAS DE UM ITEM DO SAEB EM FUNÇÃO DA PROFICIÊNCIA DOS ESTUDANTES



Exemplo de um item bom:

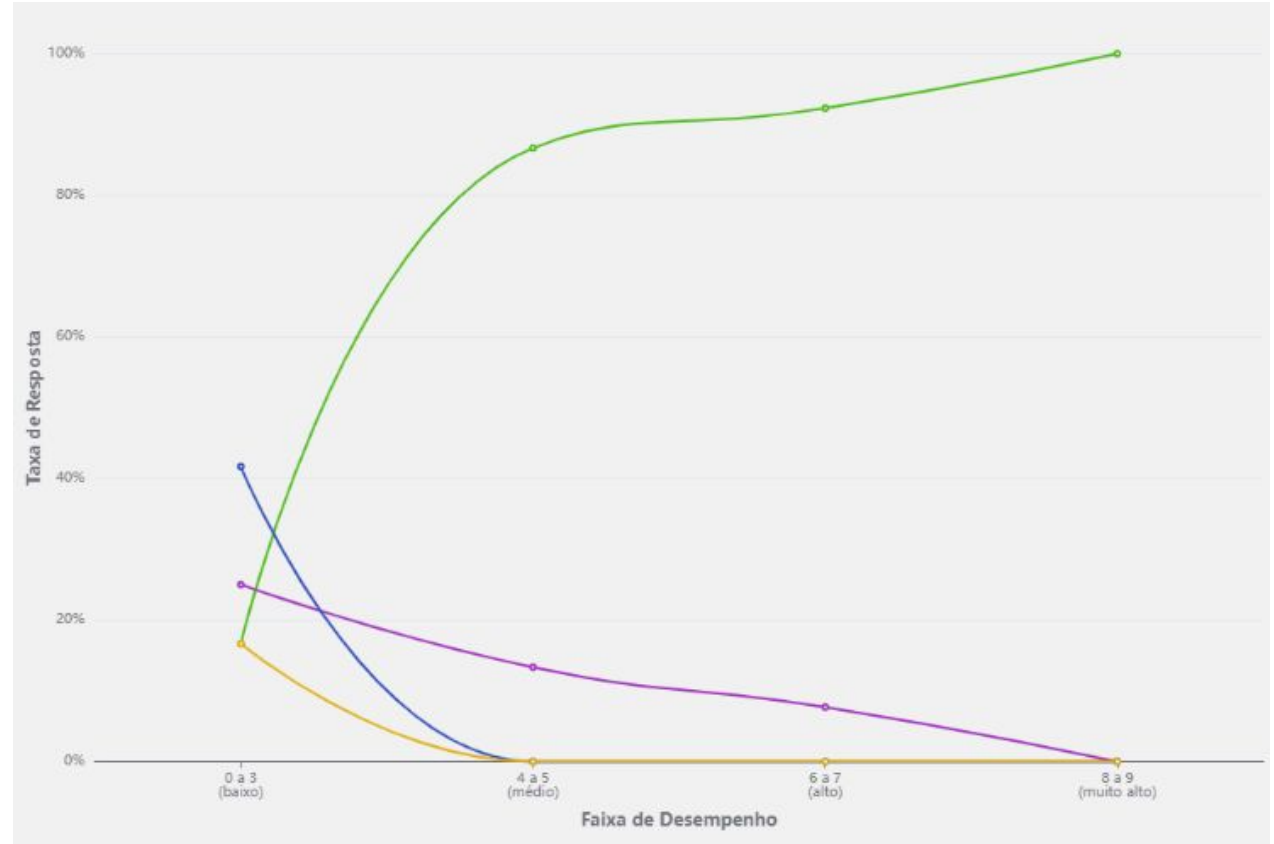
2. Gabriel encontrou seu tio em Paris. Em 2023, a soma das idades dos dois era 117 anos. E Gabriel é vinte e sete anos mais novo que seu tio. Qual é o sistema que representa esse problema?

A) $\begin{cases} x + y = 117 \\ x - y = 27 \end{cases}$ ✓

C) $\begin{cases} x + y = 27 \\ x = y + 117 \end{cases}$

B) $\begin{cases} 27x + y = 117 \\ x - y = 27 \end{cases}$

D) $\begin{cases} x + y = 27 \\ x - 27y = 117 \end{cases}$

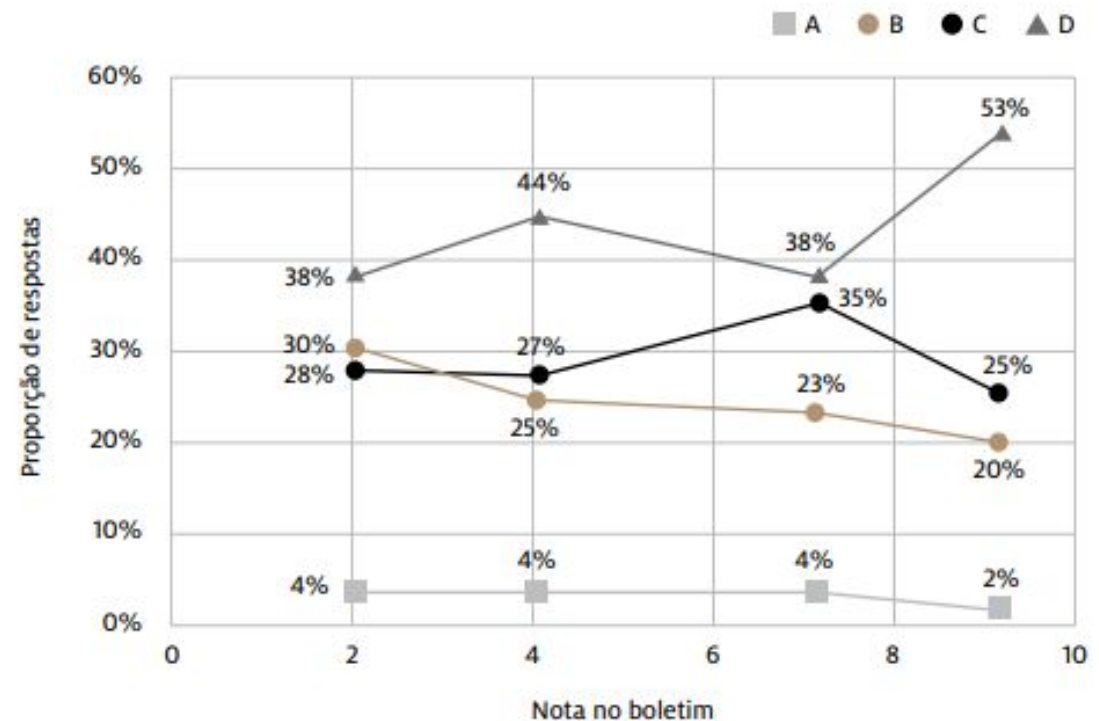


Alternativas: **A B C D**

Item ruim

- Já no **Item ruim**, os estudantes que mais acertam uma questão não necessariamente são aqueles de desempenho mais alto.
- O gráfico ao lado traz um exemplo de como se comporta um item ruim. Nesse caso, a alternativa B é a correta.
- Observe que há um percentual mais elevado de acerto (30%) entre os que têm nota mais baixa no boletim (perto de 0).

EXEMPLO DE COMO PODE SE COMPORTAR UM ITEM RUIM

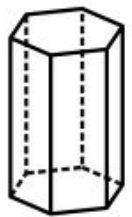


Exemplo de um item ruim:

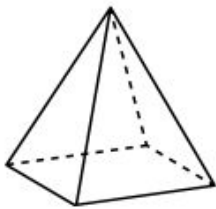
1) OBSERVE A IMAGEM ABAIXO.



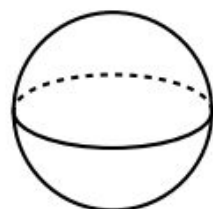
A FORMA DESTA LATA LEMBRA QUAL SÓLIDO?



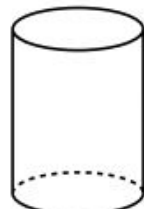
A)



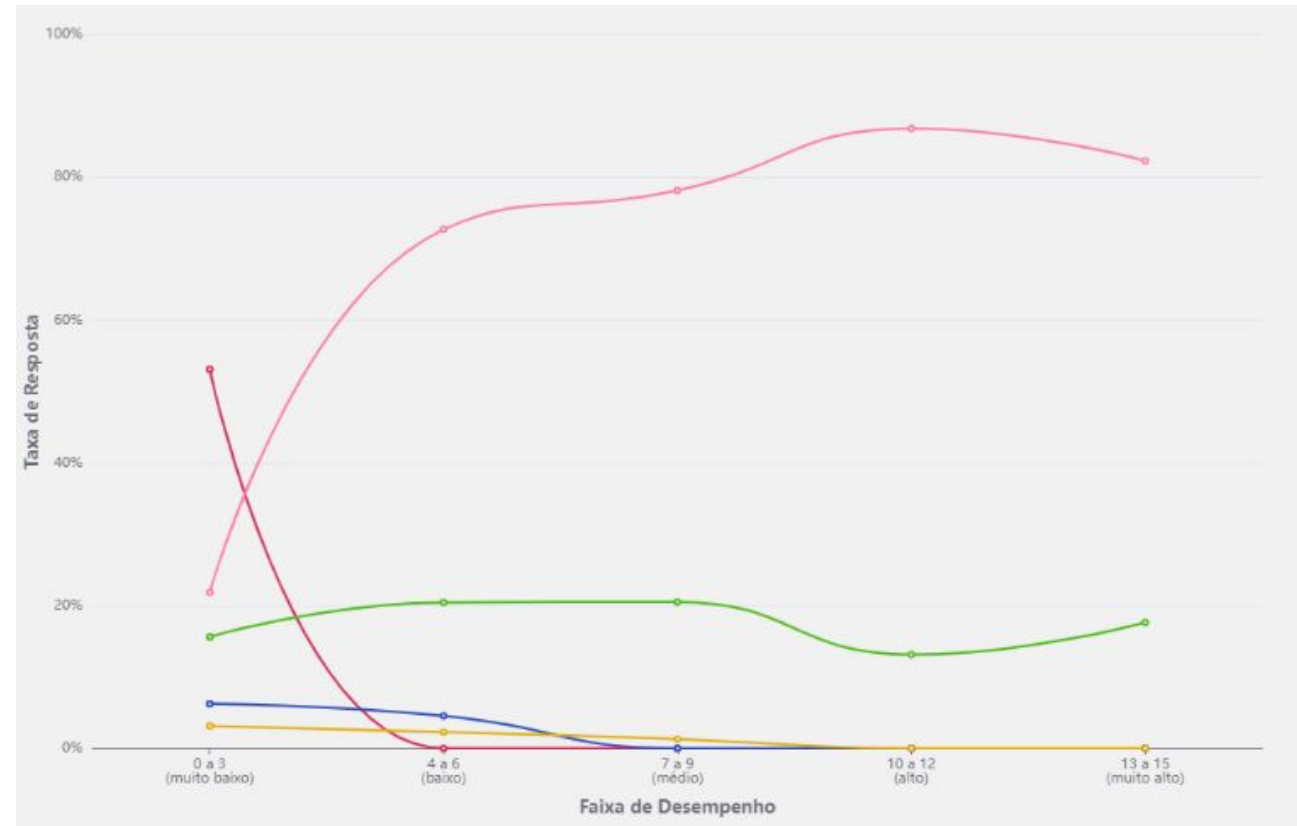
B)



C)



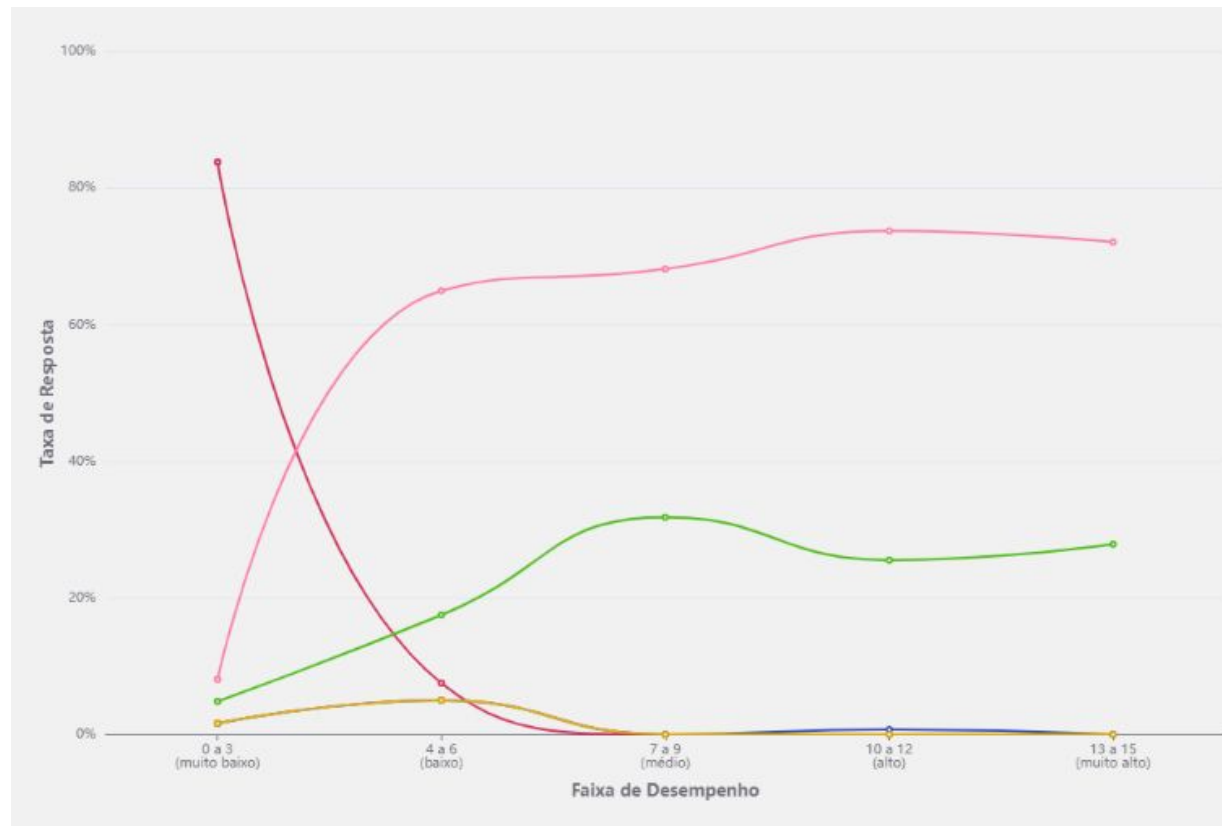
D)



Alternativas: - A B C D


Exemplo de um item ruim:

Mesma questão, respondida por outros alunos:



Alternativas: - A B C D

Exemplo de um item ruim:

 LEIA O TEXTO:

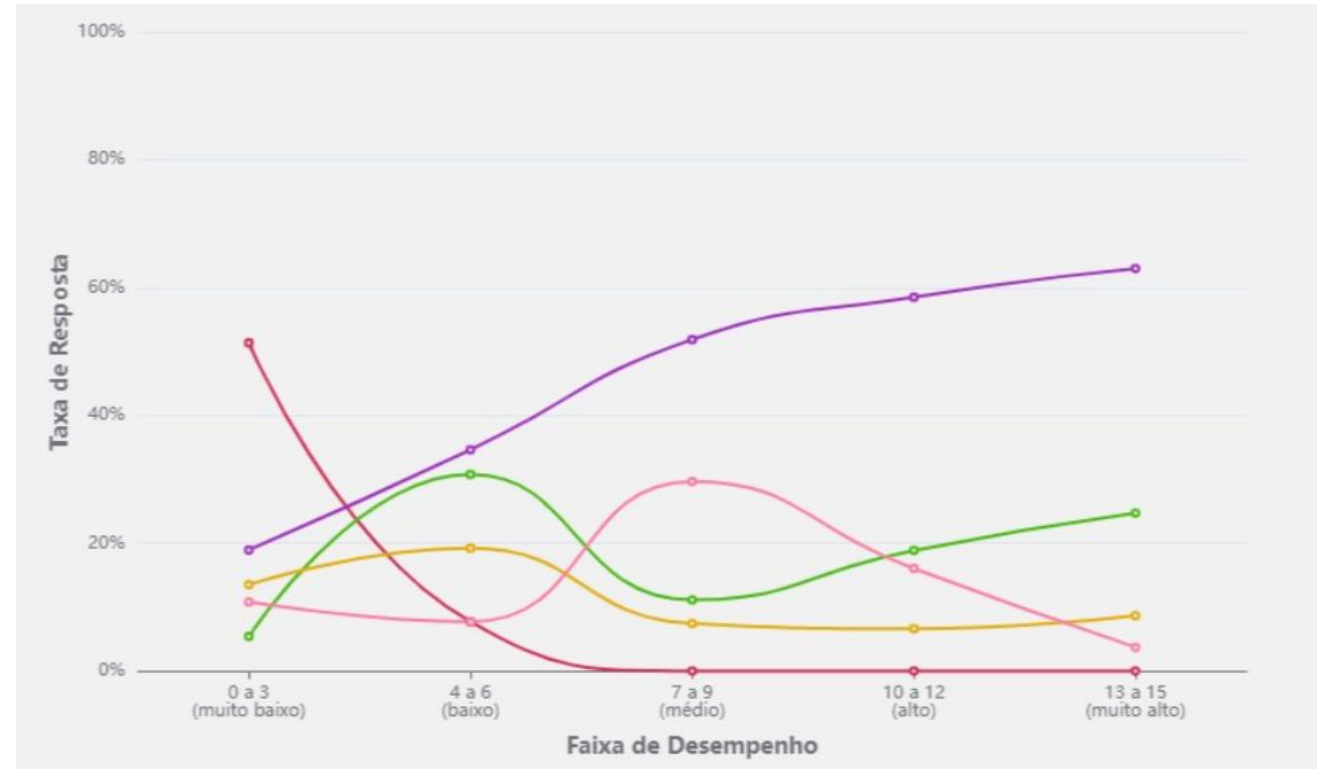
MAMÃE,

APÓS A AULA DE INGLÊS, EU E O MARCOS IREMOS PASSAR NA CASA DA TIA PAULA, POIS VAMOS PEGAR EMPRESTADO O JOGO DA MEMÓRIA DO PRIMO GABRIEL. NÃO IREMOS DEMORAR!

UM BEIJO, CARLA.

 A FRASE: NÃO IREMOS DEMORAR!

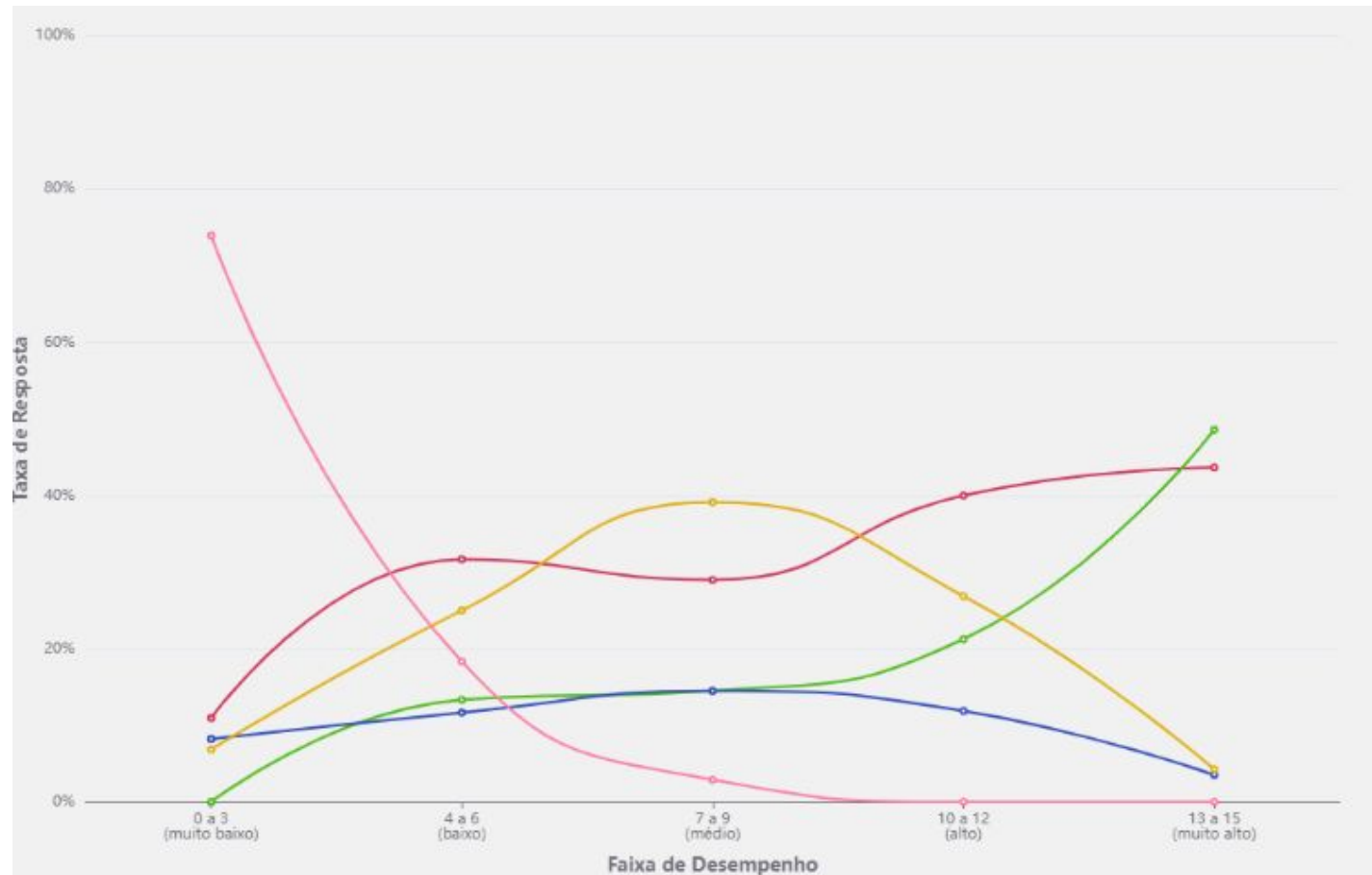
- A) É UM FATO.
- B) É UM OPINIÃO. ✓
- C) É UM DESEJO..
- D) É UMA BRINCADEIRA.



Alternativas: - A B C D


Exemplo de um item ruim:

Mesma questão, respondida por outros alunos:




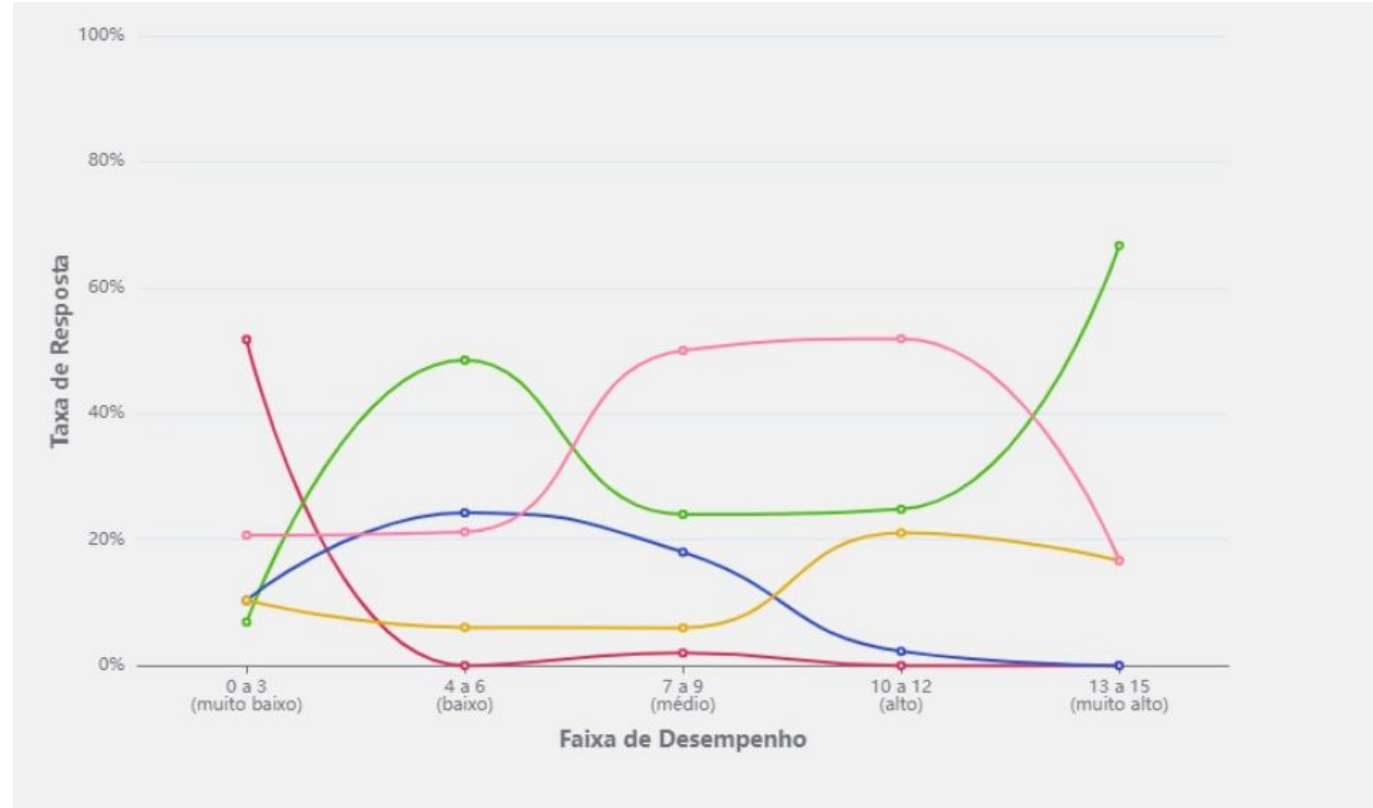
Alternativas: A B C D Não Respondeu

Outros exemplos:

 A SALADA DE SUZANA
SALADA É UM DOS PRATOS FAVORITOS DE SUZANA.
A MÃE DELA FAZ SALADA COM TOMATE, ALFACE, AZEITONA,
RÚCULA E PEPINO.
ELA SEMPRE COME TUDO.
SUZANA TEMPERA A SALADA COM SAL, AZEITE E LIMÃO E
FICA UMA DELÍCIA!
SUZANA NÃO É BOBA! ELA SABE QUE COMER SALADA FAZ
BEM PARA A SAÚDE.
E VOCÊ, COME SALADA TAMBÉM?

 NESSE TEXTO, A PARTE QUE MOSTRA UMA OPINIÃO É

- A) SALADA É UM DOS PRATOS FAVORITOS DE SUZANA 
- B) ELA SEMPRE COME TUDO
- C) E FICA UMA DELÍCIA!
- D) ELA SABE QUE COMER SALADA FAZ BEM



Alternativas: - A B C D

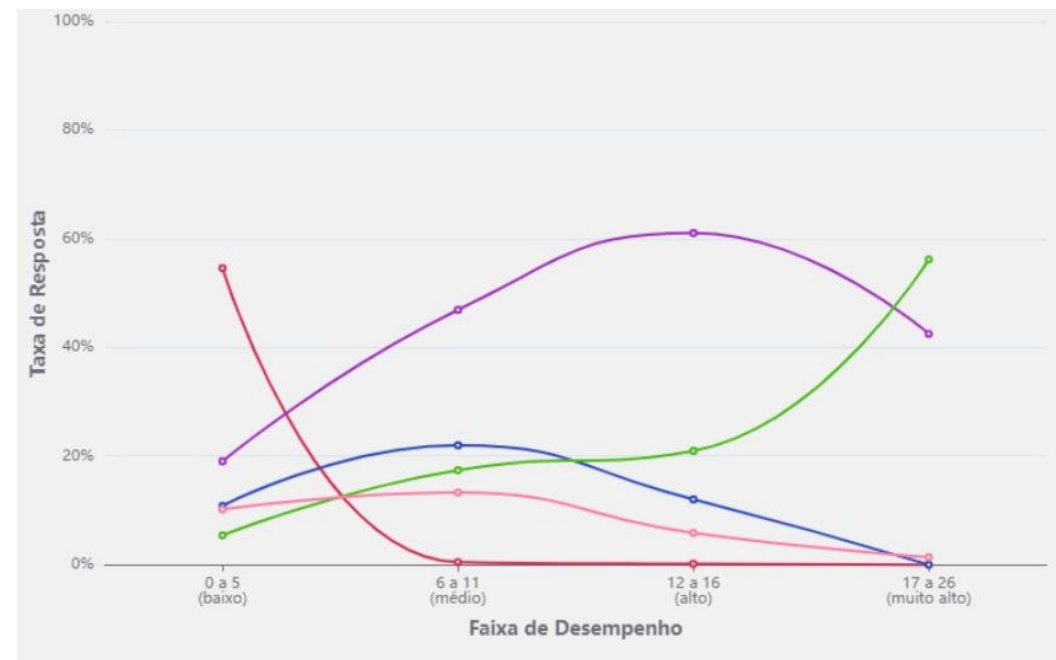
Outros exemplos:

11) (M091541H6) Considere a sequência numérica apresentada no quadro abaixo.

Sequência	11	18	25	32	39...
Posição (n)	1	2	3	4	5...

Uma expressão algébrica que modela essa sequência numérica, em função da posição n que cada termo dessa sequência ocupa, está representada em

- A) $n + 7$.
- B) $n + 11$.
- C) $7n + 4$. ✓
- D) $7n + 11$.



Alternativas: - A B C D

Outros exemplos:

Leia o texto abaixo.

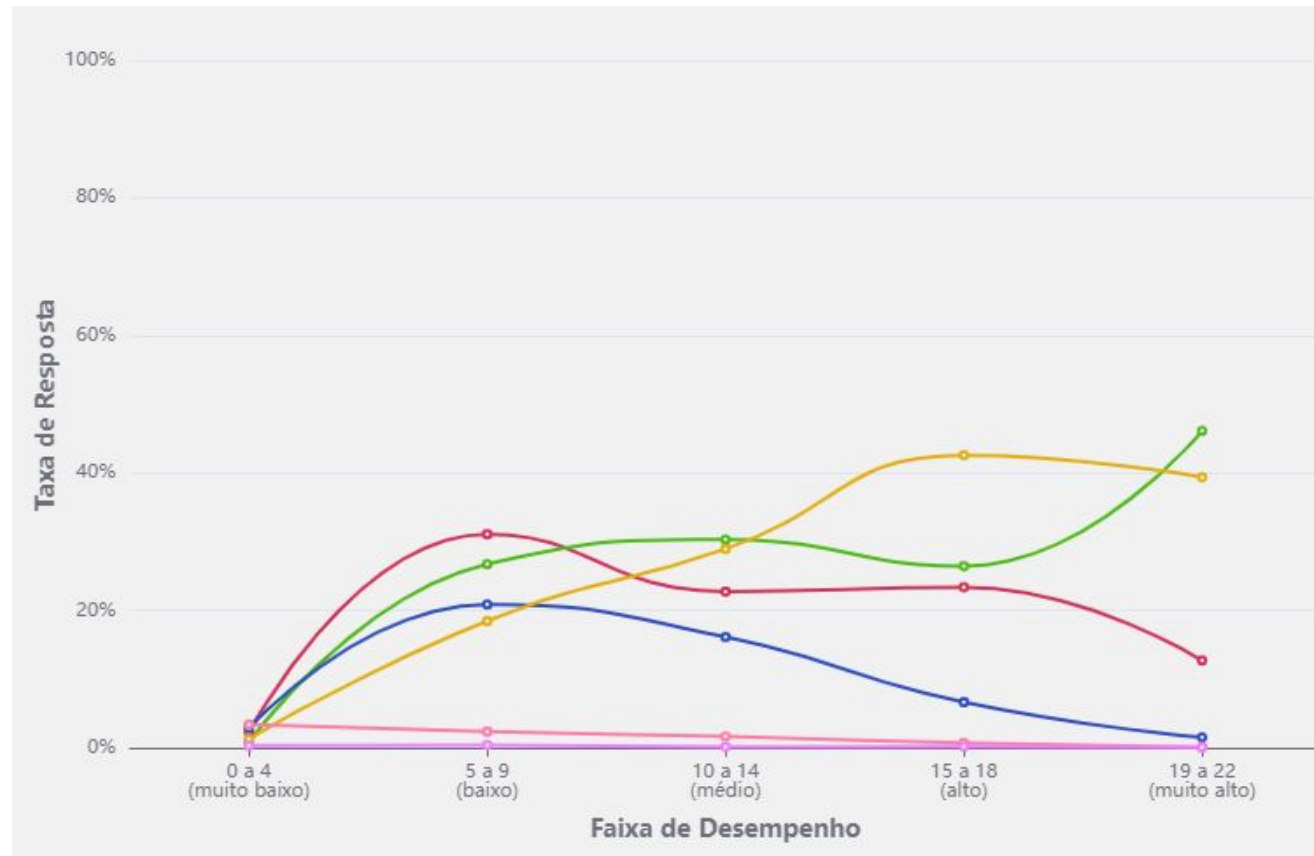
Alexandria, 22 de julho de 2004.
Oi, Samy!
Como vai você?
Estou há poucos dias no Egito, mas já começo a sentir saudades do Brasil.
Aqui está tudo bem. A família de Rameses é muito legal [...]. Ainda bem que os tios dele falam português, senão eu estaria perdida, pois não falo nenhuma palavra em árabe, a língua que se fala aqui no Egito. Alexandria é uma cidade muito linda. [...]
[...] Muitas coisas são parecidas com as nossas: eles veem televisão e assistem até a um desenho animado muito engraçado, que se chama Sipi Bad. Eu não entendo nada! Só sei que parece com Aladim [...].
E você, o que tem feito de bom? [...]

Um abraço.
Rita.

Disponível em: <<https://ibityl.com/ryLsEW/>>, Acesso em: 24 maio 2022. Fragmento. (SUP03081217)

Este texto serve para

- transmitir um ensinamento para Samy. ✓
- comunicar as experiências pessoais para Samy.
- ensinar uma brincadeira.
- enviar uma notícia.



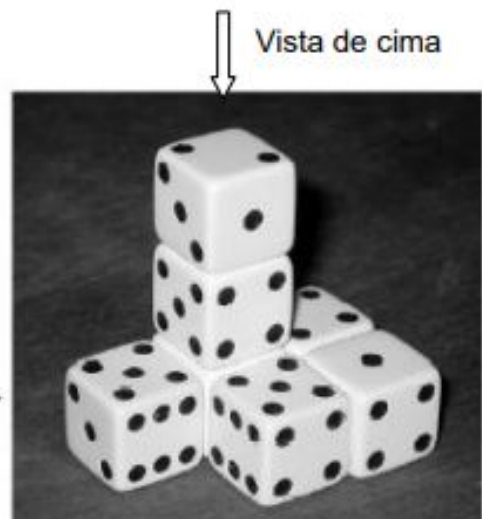
Distratores

- São opções de resposta incorretas ou enganosas em testes ou questionários.
- São projetados para refletir concepções errôneas comuns e interpretações equivocadas dos estudantes em relação a um tópico, além de ajudar a compreender até onde vai o nível de conhecimento dos alunos para esse determinado tópico.
- São utilizados para avaliar a compreensão e o conhecimento dos estudantes em áreas específicas do conhecimento.
- Podem indicar domínios parciais dos alunos em relação a determinados temas.
- Desafiam os estudantes a identificar e selecionar a resposta correta entre opções incorretas atrativas ou plausíveis.



UMA CONSTRUÇÃO COM DADOS

A figura abaixo mostra uma construção feita com sete dados idênticos, cujas faces estão numeradas de 1 a 6.



Quando a construção é olhada de cima, somente 5 dados podem ser vistos.

Questão 1: UMA CONSTRUÇÃO COM DADOS

PM937Q01 - 0 1 2 9

Quantos pontos ao todo podem ser vistos, quando esta construção é olhada de cima?

Número de pontos vistos:

Resposta correta; 17

Distrator: 16



Apresentação da Plataforma