

Análise Clássica dos Testes
(TCT)

METODOLOGIA

A primeira análise apresentada no relatório é uma informação resumo em relação ao número de respondentes, número de itens e medidas de desempenho dos respondentes.

Medida de qualidade da prova/instrumento

Na sequência a confiabilidade do instrumento/prova é analisada por meio da consistência interna do instrumento e para isso é utilizado o coeficiente Ômega (Ω) de McDonald, que é uma medida de correlação do item com o traço latente que está sendo medido pelo instrumento/prova. Ao analisar o traço latente, espera-se que o instrumento tenha um traço latente dominante, como por exemplo, conhecimento em matemática, ou seja, todos os itens estariam relacionados com a proficiência dessa área do conhecimento. Um instrumento com boa confiabilidade deve ter $\hat{\Omega} \geq 0,70$.

Além do Ômega geral do instrumento, é apresentado também o Ômega com a exclusão de cada item do instrumento/prova. Neste caso, considerando a exclusão do item, se o valor do Ômega passar a ser maior do que o Ômega geral, esse item pode não estar diretamente relacionado ao traço latente do instrumento/prova e é colocado na frente do valor do item uma interrogação "?", indicando que esse item deve ser observado com mais cuidado.

Um exemplo da saída da informação do Ômega de uma prova com 13 itens é apresentado a seguir.

ÍNDICE DE CONFIABILIDADE
Ômega= 0,967

Item Excluído	Ômega	
I1	0,967	.
I2	0,965	.
I3	0,963	.
I4	0,964	.
I5	0,963	.
I6	0,963	.
I7	0,964	.
I8	0,966	.
I9	0,964	.
I10	0,966	.
I11	0,969	?
I12	0,965	.
I13	0,964	.

O cálculo desse índice considera uma medida de desempenho dos respondentes do instrumento/prova, dada pelo Escore (número de respostas corretas).

Medidas associadas aos itens

O **Índice de Dificuldade (DIF)** é representado na tabela como uma porcentagem (PORC), que varia de 0% a 100%. Para itens de múltipla escolha, um valor alto (95%, digamos) para a categoria gabarito significa que o item é fácil para o grupo analisado, enquanto um valor baixo (25%, digamos) significa que o item é difícil. Para itens politômicos ordinais (itens discursivos que apresentam pontuações parciais), um valor alto na categoria mais alta (mais correta) indica que o item é fácil, caso contrário, ou seja, um valor baixo na categoria mais alta (mais correta), mostra que o item é difícil.

A **Correlação Bisserial (BIS)** é uma medida de diferenciação do item e varia de -1 a 1. Para itens de múltipla escolha, um valor da Correlação Bisserial negativo na categoria gabarito é um indicativo de item que não apresentou um bom funcionamento para a população analisada, pois aponta que respondentes de baixo escore devem estar marcando mais a categoria

gabarito do que os respondentes de escore mais alto, uma situação não apropriada. Também são apresentadas estatísticas dos distratores do item, sendo desejáveis valores negativos nessas categorias. Para itens politômicos, espera-se que a categoria de resposta mais alta (maior acerto) apresente o maior valor (e positivo) de Correlação Bisserial e, gradativamente, para as categorias mais baixas, o valor da Correlação Bisserial deve diminuir até a categoria mais baixa, para a qual se espera que a Correlação Bisserial seja negativa.

O **Índice de Discriminação** (DISC) analisa a porcentagem de respostas em dois grupos. O grupo inferior (INF) é formado pelos respondentes com 27% dos menores escores e o grupo superior (SUP) pelos respondentes com 27% dos maiores escores. O índice varia entre -1 e 1, sendo calculado como a diferença entre as porcentagens de respostas apresentadas pelo grupo superior e pelo grupo inferior. Para itens de múltipla escolha espera-se um valor positivo para a categoria gabarito, indicando que o grupo superior tem maior quantidade de respostas na categoria do gabarito do que o grupo inferior, o que é desejável para um item de qualidade. Para itens politômicos, espera-se que a categoria mais alta (maior acerto) apresente o maior valor (e positivo) do índice e, gradativamente, para as categorias mais baixas, o valor do índice deverá reduzir até a categoria mais baixa, quando é esperado que o Índice de Discriminação seja negativo.

Além da informação individual de cada item, também é apresentada uma informação resumo de todos os itens em relação ao gabarito. A fim de obter essa medida resumo da qualidade dos itens, utilizou-se os critérios de classificação dos itens estabelecidos a seguir. É importante reforçar que para os itens de múltipla escolha essa classificação se dá em relação ao gabarito de cada item; e para os itens politômicos essa classificação é para a categoria mais alta (maior acerto).

Índice de Dificuldade (DIF) do gabarito/categoria mais alta

Dificuldade	Mínimo	Máximo
Muito difícil	0%	< 15%
Difícil	15%	< 35%
Médio	35%	< 65%
Fácil	65%	< 85%
Muito fácil	85%	100%

Coefficiente Bisserial (BIS) do gabarito/categoria mais alta

Bisserial	Mínimo	Máximo
Inapropriado	-1,000	< 0,000
Inadequado	0,000	< 0,100
Moderado	0,100	< 0,200
Adequado	0,200	< 0,300
Excelente	0,300	1,000

Índice de Discriminação (DISC) do gabarito/categoria mais alta

Discriminação	Mínimo	Máximo
Não discrimina	-1,000	< 0,000
Baixa	0,000	< 0,100
Moderada	0,100	< 0,200
Adequada	0,200	< 0,300
Excelente	0,300	1,000

A classificação de cada item, em relação ao seu gabarito (ou categoria mais alta para itens politômicos), é colocada em uma tabela no relatório e ao final da tabela é acrescentada uma coluna com uma “Indicação” de revisão do item, caso o item apresente em seu gabarito $BIS < 0,1$ ou $DISC < 0,1$.

Algumas tabelas e gráficos resumo com a informação da quantidade de itens em cada classificação, bem como o cruzamento da classificação entre índice de dificuldade, coeficiente bisserial e índice de discriminação são também realizados, a fim de analisar o mapeamento dos itens do instrumento/prova nos três índices de qualidade. Nesta análise não se deseja encontrar itens nas classificações “Inapropriado” e “Inadequado” do coeficiente bisserial; e nas classificações “Não discrimina” e “Baixa” do índice de discriminação.

Análise gráfica dos itens

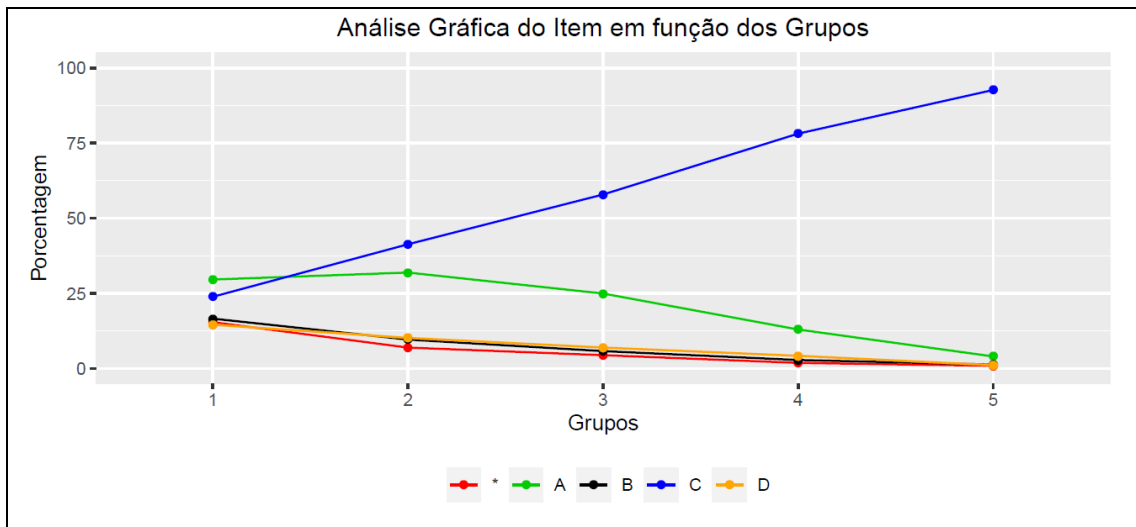
Uma análise gráfica também é realizada para mostrar as porcentagens de resposta para cada alternativa por segmentos consecutivos dos respondentes, definidos pelos grupos de escore. Se o total de respondentes é menor ou igual a 300 são definidos 3 grupos; se o total de respostas for entre 301 e 1.000 são definidos 5 grupos; e se o total de respostas for maior que 1.000 são definidos 7 grupos. A seguir é exemplificada a divisão dos grupos, considerando o caso com 5 grupos.

Grupos	Respondentes ordenados pelo Escore
1	0% a <20%
2	20% a <40%
3	40% a <60%
4	60% a <80%
5	80% a 100%

Para itens de múltipla escolha, deve-se observar no gráfico se a linha relativa à alternativa do gabarito é crescente, o que significa que grupos de escores mais altos tendem a ter maior porcentagem de acertos do item em análise. Por outro lado, espera-se que a linha relativa as alternativas dos distratores seja decrescente.

Para o gráfico descrito acima, nos itens politômicos espera-se que a linha mais crescente seja da categoria mais alta (maior acerto), mas o item também pode apresentar uma linha crescente para outras categorias de acertos parciais, porém para as categorias mais baixas, as linhas devem ser decrescentes.

A tabela e gráfico de cada item, contidos nos relatórios, apresentam o resumo estatístico do item, conforme exemplo abaixo. As informações apresentadas para cada item poderão ser analisadas individualmente pelo especialista que elaborou o item. A categoria gabarito é sempre representada na cor azul. O exemplo que segue é de um item de múltipla escolha.



Índices da TCT

	ALT	PORC	BIS	DISC	INF	SUP	Gab
A		20,7	-0,271	-0,206	29,1	8,5	.
B		7,2	-0,358	-0,126	14,7	2,1	.
C		58,7	0,551	0,551	30,2	85,3	<=
D		7,4	-0,288	-0,102	12,8	2,6	.
*		5,9	-0,402	-0,117	13,2	1,5	.

RESULTADOS
DA
ANÁLISE