PRODUTO EDUCACIONAL

MANUAL DO SOFTWARE XCALIBRE PARA APLICAÇÃO DO MODELO DE RASCH

JOSÉ MARIA FERREIRA DE FARIAS

BRASÍLIA
2015
JOSÉ MARIA FERREIRA DE FARIAS

Este manual faz parte da dissertação: *Método Para Otimizar o Aprendizado no Ensino Médio com base no Modelo de Rasch* – requisito parcial à obtenção do Título de Mestre em Ensino de Física, na Área de Concentração “Física na Educação Básica”, pelo Programa MNPEF - Mestrado Nacional Profissional em Ensino de Física, realizado pela Universidade de Brasília e Sociedade Brasileira de Física. Orientação do Prof. Dr. Joaquim José Soares Neto,

BRASÍLIA
2015
Este manual simplificado do software Xcalibre é uma porta aberta àqueles professores e avaliadores, interessados na Teoria de Resposta ao Item (TRI), que pretendem mergulhar numa visão moderna de avaliação para descobrir o quanto e o quê o aluno aprendeu.
Prezado professor,

As novas possibilidades de avaliação, propiciadas pelo avanço das tecnologias da ciência informacional, oferecem maior precisão do que os métodos até então utilizados, que se baseiam na Teoria Clássica dos Testes (TCT). Esta dá apenas escore bruto, sem discriminar o que o aluno aprendeu.

A Teoria de Resposta ao Item (TRI) fornece informações além do escore bruto, também do quanto e do quê o aluno aprendeu. Os softwares e hardwares hoje disponíveis permitem ao professor e avaliador o uso desta moderna ferramenta de avaliação.

O SAEB, Sistema de Avaliação do Ensino Básico, e o ENEM, Exame Nacional do Ensino Médio, fazem uso da Teoria de Resposta ao Item (TRI), ao invés da Teoria Clássica dos Testes (TCT), para promover a seleção de alunos em função da sua habilidade de dar a resposta correta sobre determinado tema.

A utilização do software Xcalibre facilita ao usuário experimentar os diversos modelos da TRI para fazer essas análises. No nosso caso, usamos como exemplo o modelo de um parâmetro, referido como modelo de Rasch, usado no exame do SAEB. Para usar o modelo de três parâmetros, empregado pelo ENEM, se faz necessária apenas uma escolha de função do software, ou seja, é bastante clicar em uma outra janelinha. Assim podem ser obtidos parâmetros avaliativos essenciais como a dificuldade do item, sua discriminação e o acerto ao acaso. As interpretações dos dados obtidos cabem ao autor da pesquisa e sua equipe.
INTRODUÇÃO

Os modelos para obtenção dos parâmetros de dificuldade dos itens de um teste/avaliação têm sido refinados pelas teorias psicométricas com o passar dos anos. Inicialmente, pela Teoria Clássica dos Testes (TCT), a proficiência do aluno ou a dificuldade de um item era simplesmente a proporção entre a quantidade de acertos e erros cometidos pelos participantes. Esse índice de proficiência é determinado pela seguinte equação:

\[ S(\%) = \frac{N(A)}{N} \times 100 \]

S = escore bruto em percentual

N(A) = número de itens acertados

N = número total de itens

Esse modelo seguido pela TCT trata apenas de um nível percentual, e por isso nada diz a respeito dos respondentes e dos itens como um todo. Com isso, os psicometristas começaram a perceber que tais análises não seriam realmente profundadas para qualificar o que o aluno aprendeu.

A TRI

A Teoria da Resposta ao Item (TRI) é um conjunto de modelos matemáticos que relacionam um ou mais traços latentes (não observados) de um indivíduo com a probabilidade deste dar uma certa resposta a um item. (ANDRADE, 2001)

A TRI se baseia na ideia de que, independentemente do indivíduo, podemos construir uma escala para medir a probabilidade da sua resposta correta a uma pergunta. Ela é composta de um conjunto de modelos matemáticos adequados a tipos determinados de itens. Graças a sua capacidade de discriminar e avaliar, ela é utilizada no Brasil pelo Sistema de Avaliação da Educação Básica (Saeb), que usa a escala de um parâmetro, a
Escala de Rasch, e pelo ENEM (Exame Nacional do Ensino Médio), que utilizam a escala de três parâmetros proposta por Lord.

No mundo, outros institutos utilizam a TRI, como o Pisa (Programa Internacional de Avaliação do Estudante, Programme for International Student Assessment), o TOEFL (Teste de Inglês como Língua Estrangeira, Test of English as a Foreign Language), o GRE (Graduate Record Examination), dentre outros.

**MODELO DE RASCH**

O modelo de Rasch é um modelo particular da TRI, de um parâmetro, usada para medir o traço latente, como no caso da qualificação profissional de um médico, da satisfação de um cliente, da existência do distúrbio alimentar, da reabilitação de um paciente, da proficiência no âmbito escolar, por exemplo, o que mostra o grande potencial deste princípio fundamental de medição em pesquisa com humanos. (BOND, 2006)

Segundo Pasquali, usando um questionário criteriosamente elaborado, a Escala de Rasch pode ser empregada para medir a proficiência do aluno e a dificuldade apresentada em cada questão. Desse modo se pode identificar a parte do programa que deve ser estimulada para superação das deficiências apresentadas.

A partir da posse de valores que representam as habilidades do aluno e as dificuldades de cada item, especialistas curriculares podem sugerir a seleção e o ordenamento adequado dos conteúdos programáticos, com a observação do tempo disponível do professor com o aluno em sala de aula.

O modelo matemático de Rasch (modelo dicotômico de um parâmetro) mede a probabilidade de o respondente dar a resposta correta a um item, e é representado a seguir:
Neste caso, o parâmetro é a dificuldade $b$.

Podemos contar ainda com outros dois modelos dicotônicos, de dois ou três parâmetros.

Dois parâmetros – a dificuldade $b$ e a discriminação $a$

Três parâmetros - a dificuldade $b$, a discriminação $a$ e o parâmetro que corrige o acerto ao acaso, $c$.

**OS PARÂMETROS**

O parâmetro $b$
O parâmetro \( b \) informa o grau de dificuldade que o item oferece ao respondente. Ele é a chave para medir o traço latente. Na Curva Característica do Item, lê-se o valor desse parâmetro na mesma escala das habilidades, quando a probabilidade de acerto do indivíduo é de 50% (cinquenta por cento).

**O parâmetro \( a \)**

O parâmetro \( a \) identifica o poder que a questão tem em discriminar a habilidade dos estudantes. Ele expressa o nível de informação do item.

**O parâmetro \( c \)**

Já o parâmetro \( c \) é utilizado para corrigir o resultado no caso de acerto do item difícil pelo respondente com pouca habilidade. (Ele corrige o chute!)

A TRI mede o traço latente, que é a habilidade não diretamente observável. Essa habilidade pode ser estimada usando qualquer uma das três escalas apresentadas, que levam em conta as características das questões, sua dificuldade, e a probabilidade de o indivíduo dar a resposta correta ao item conforme a sua habilidade.

**Unidimensionalidade**

Para aplicação do método de Rasch pressupõe-se:

A *unidimensionalidade*, pois o conjunto de itens mede apenas uma dimensão, uma variável latente. Por isso, se eliminam os itens que violam este princípio.

**Independência Local**

A *independência local*, ou seja, nenhum item deve conter informação que possa ser usada para responder a um outro. Então, devido à inexistência de correlação entre os itens, a probabilidade de um sujeito acertar \( N \) itens é igual ao produto das probabilidades de ele acertar cada um deles. – Pasquali salienta
a dificuldade de se estabelecer a independência local, mas se os fatores estranhos forem controlados, mantidos constantes, o fator dominante, a dificuldade do item, será a única fonte de variação. Neste caso, as respostas se tornam independentes, já que o examinando conta apenas com a magnitude da sua habilidade para responder às questões.

Calibragem

A calibragem do item significa equiparar, tornar comparáveis os parâmetros dos itens provenientes de testes diferentes e os traços latentes de respondentes de diferentes grupos na mesma métrica, tornando assim os itens e os respondentes comparáveis. Com um banco de itens calibrados, em que os itens e seus parâmetros estejam na mesma escala de proficiência, podemos construir um ou mais testes com graus de dificuldade que atendam aos objetivos de uma ou de mais avaliações.

SOFTWARE XCALBIRE (VERSÃO 4.1.6)

Para efetuar-se as análises de dificuldade dos itens através do Modelo Rasch (ou de 1 parâmetro), optou-se pela utilização do programa estatístico XCALIBRE (Versão 4.1.6). O principal motivo para a escolha desse programa foi a sua facilidade de manipulação, bem como a sua disponibilidade no mercado de softwares que analisam os parâmetros da TRI, além de ser encontrado em versão livre.


O que segue é um passo a passo para utilização do programa, utilizando-se de exemplos para a realização de análises de parâmetros de dificuldade do item pela Teoria de Resposta ao Item.
Um dos diferenciais da utilização do XCALIBRE é que o programa não necessita que o usuário seja um *expert* em TRI. Existem abas no programa que auxiliam na obtenção dos resultados que o usuário desejar, a partir do momento em que ele conheça o que essas abas podem fornecer.

Quais seriam essas abas? Veja na figura abaixo:

As abas acima estão numeradas e serão explicadas no trabalho posteriormente. É necessário saber apenas algumas características básicas:

1) *Files* (essa aba é onde o usuário irá inserir os arquivos de entrada, controle e saída de dados);

2) *Input Format* (é onde o usuário irá inserir informações que irão ajudar o programa a ler os bancos, como, por exemplo, a partir de que momento o programa deve ler os dados, se há dados omissos, se todos os dados serão lidos na análise, etc.);

3) *IRT Model* (nessa aba o usuário poderá escolher qual modelo deseja utilizar, a partir do tipo de itens (dicotômico, politômico ou misto), e de qual modelo gostaria de se utilizar (1, 2 ou 3 parâmetros);
4) *Calibration* (nessa aba o usuário pode selecionar distribuições *a priori* para as suas análises, bem como critérios de convergência e escalonamento de parâmetro de Item);

5) *Estimation* (quando se clica nessa aba, tem-se acesso a características relacionadas às estatísticas clássicas dos itens, e de que forma irá ser estimada a aptidão [*Theta*] do indivíduo; e

6) *Output Options* (basicamente estabelece como o arquivo de *output* [ou saída] será demonstrado para o usuário).

Veja agora a função de todas as abas, com alguns detalhes:

1 – Aba *Files*

Essa aba, que já foi mostrada na figura anterior, é o espaço no qual o usuário irá inserir três arquivos essenciais para a análise: 7) *Data Matrix File* (aqui devem ser colocados os dados brutos digitados pelo usuário*), 8) *Item Control File* (esse é o lugar onde o usuário deve inserir o “gabarito” do seu teste, com algumas outras informações que serão definidas na próxima seção), e 9) *Output File* (o usuário irá inserir aqui um arquivo no formato RTF [*Rich Text Format*] para que o programa possa fornecer as análises realizadas nesse arquivo).

Veja como é possível construir esses bancos de dados nas seções que se seguem.

1.1 – Construindo o arquivo de entrada (*Data Matrix File*)

Um arquivo de entrada deve ter algumas características essenciais para que o programa possa lê-lo. Elas são as seguintes: a) A identificação de cada respondente, e b) as respostas fornecidas por esses respondentes a cada item. Em cada linha está a informação de uma pessoa. Veja no exemplo abaixo a construção envolvendo 5 pessoas e 19 itens respondidos.

| 01 | 2245333251313423431 |
| 02 | 2315213513244154545 |
| 03 | 3322341231443341251 |
| 04 | 5555343114 24123322 |
| 05 | 1245542314232314532 |
Os itens acima possuíam cinco possibilidades de respostas (A, B, C, D e E). As respostas foram colocadas como números para ajustar e facilitar a leitura do programa, dessa forma A = 1, B = 2, C = 3, D = 4, E = 5. É importante notar também que a quantidade de espaços será importante quando for o momento da utilização da aba *Input Format*, pois erros cometidos na quantidade de espaços fará com que o programa leia características que nada tem a ver com a variável ou não lei determinada característica por não estar no local correto.

1.2 – Construindo o arquivo de controle (*Item Control File*)

A função desse arquivo de controle é inicialmente fornecer dados para que o programa possa avaliar as questões como certas ou erradas, e, posteriormente, com base no que foi analisado, estabelecer as relações entre os tipos de itens, se eles serão utilizados na análise, que tipo de domínio estão avaliando etc. A seguir, é mostrado um exemplo de um arquivo de controle construído para o exemplo acima. Note que cada característica do item será explicada também.

<table>
<thead>
<tr>
<th>Item 1</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>Y</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>Y</td>
<td>M</td>
</tr>
<tr>
<td>Item 3</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>Y</td>
<td>M</td>
</tr>
<tr>
<td>Item 4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>Y</td>
<td>M</td>
</tr>
<tr>
<td>Item 5</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>Y</td>
<td>M</td>
</tr>
</tbody>
</table>

**Figura 2. Exemplo de arquivo de controle construído para o exemplo da Figura 1.**

As letras em vermelho possuem o seguinte significado:

A ➔ Essa coluna identifica para o programa qual item está sendo avaliado. É possível atribuir qualquer nome para as informações nessa coluna, porém o nome Item facilita a interpretação posterior no arquivo de saída. (Apenas
duas exceções: vírgulas e espaços usando a tecla *Tab* [encontrada logo acima da tecla *Caps Lock* ou “Fixa”]

**B ➔** Na segunda coluna, identifica-se a resposta correta do item. No caso, os itens variaram de 1 até 5. Caso os itens sejam dicotônicos, ou seja, certo ou errado, devo inserir apenas o número 1 para todas as opções. (Importante: podem ser usadas letras, caso sejam de múltipla-escolha, e se forem politônicos atribuí-se o sinal “+” se forem positivamente avaliados, e “-” se forem negativamente avaliados).

**C ➔** Esse é o número de alternativas do item. (O programa permite até um número limite de 15 alternativas por item. Para itens dicotônicos, avaliados como 0 e 1, erro ou acerto, o número de alternativas seria 2);

**D ➔** A coluna D mostra a área de domínio ou conteúdo (que pode variar de 1 até 50). Essa coluna possui uma análise mais subjetiva, e é necessário identificar bem o que cada item avalia para saber se ele faz parte do mesmo domínio dos outros.

**E ➔** Nessa coluna, mostra-se o status de inclusão de um item. As legendas que se seguem estão em inglês:

1) **Y** = Sim (ele foi incluído)
2) **N** = Não (o item não foi incluído)
3) **P** = O item fez parte do pré-teste
4) **A** = Um item âncora, que apresenta parâmetros já identificados e que podem influenciar os resultados das análises a partir das suas inserções ou ausências.

**F ➔** O tipo de item do teste. São especificados três tipos de item para o teste:

1) **M** = Itens de múltipla-escolha ainda não avaliados, com respostas que começam com 1 ou A
2) **R** = Itens politônicos que começam com 1 ou A
3) **P** = Itens com respostas numéricas que começam em 0 (p.ex., 0, 1, 2 e 3). Inclui-se aqui itens de crédito parcial de categoria múltipla, e itens de múltipla-escolha avaliados dicotomicamente.

Importante notar que cada informação fornecida nesse arquivo de controle é de total responsabilidade do usuário, e que os erros realizados nessa construção irão fornecer resultados errados quando da análise.
1.3 – Construindo o arquivo de saída (ou *Output*)


2 – A aba *Input Format*

Nessa aba, o usuário identifica características relativas ao arquivo de entrada de dados. Note a seguir como é essa tabela e algumas das suas características.

![Input Format](image)

**Figura 2.1 – Opções da aba *Input Format***

A **Number of examinee ID columns**: Para explicar melhor essa opção, deve-se retornar ao arquivo de entrada para avaliar as possibilidades. Em tradução
livre para o português, essa opção seria “Número de colunas para identificação do examinando”. Note as colunas na figura a seguir.

<table>
<thead>
<tr>
<th>Pers01</th>
<th>100011100044345</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pers02</td>
<td>111100010101000</td>
</tr>
<tr>
<td>Pers03</td>
<td>100414115114100</td>
</tr>
<tr>
<td>Pers04</td>
<td>111100010101100</td>
</tr>
</tbody>
</table>

Figura 2.2 – Exemplo de dados para o arquivo de input

Note que a identificação do examinando possui seis colunas de caracteres (p.ex.: pers01 – 4 letras e 2 números = 6 caracteres), logo é essa informação que o programa entenderá (“Qualquer informação acerca da identificação irá somente até a coluna 6”).

B ➔ Examinee IDs begin in column: Essa opção fornece informação acerca de qual coluna inicia a identificação do examinando. No caso do exemplo, qualquer dado que estiver na coluna 1 já fará parte da identificação do indivíduo. Assim, as opções A e B delimitam qual o início e o final da informação relativa ao respondente. (“A coluna 1 fornece informação inicial sobre o respondente.”). (Importante: o programa admite até 999 caracteres para essa identificação.)

C ➔ Item responses begin in column: A opção marcada com a letra C serve para dizer se há espaços ou não entre as informações do respondente e as respostas emitidas por ele. É importante que se tenha pelo menos um espaço para separar onde terminam as informações do respondente e onde começam as informações dos itens. (Um número 7 nessa opção diz que a partir da sétima coluna começam as respostas do respondente.)

D ➔ The data matrix file is delimited by a: Essa opção é facultativa e deve ser marcada quando:

a) os itens forem separados por vírgula (comma) e
b) os itens forem separados por uma tabulação (tab).

Segue um exemplo de cada arquivo.

Pers01,1,0,0,0,1,1,1,0,0,0,4,4,3,4,5
Pers02,1,1,1,1,0,0,0,1,0,1,0,1,0,0,0
Pers03,1,0,0,4,1,4,1,1,5,1,1,4,1,0,0
Pers04,1,1,1,1,0,0,0,1,0,1,0,1,1,0,0

Figura 2.3 – Exemplo de arquivo de entrada delimitado por vírgulas
Note que entre a coluna final com informações relativas à identificação do respondente e a coluna inicial com informações relativas às respostas do item não há espaçamento, apenas vírgula delimitando cada um delas.

<table>
<thead>
<tr>
<th>Pers01</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>4</th>
<th>4</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pers02</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pers03</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Pers04</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Figura 2.4 – Exemplo de arquivo de entrada limitado por tabulação

No exemplo da figura 2.4 há apenas o espaço de uma tecla tab entre cada item e entre o primeiro item e a última coluna de informação do respondente.

Mas qual seria melhor? Ambas são lidas da mesma forma pelo programa, essas seriam apenas questões de estilo na construção do arquivo.

É importante notar que caso o usuário selecione a opção Delimited Data, a opção Fixed width Data não estará selecionável. Isso ocorre pois o programa consegue estabelecer, pela opção Delimited Data, que as vírgulas ou as tabulações são marcadores de início e fim de uma determinada categoria (a saber, informações dos respondentes e informações dos itens).

E ➔ Omit Character: Nessa opção é possível atribuir qualquer caractere para um determinado item que será omitido na análise que será realizada. Considera-se um item omitido aquele que o sujeito deixou de fazer mesmo tendo feito o imediatamente anterior e o posterior a ele. (IMPORTANTE: Para itens dicotômicos, um item omitido será tratado como chute do examinando, no modelo de 3 parâmetros. Enquanto no de 1 e 2 parâmetros será tratado apenas como erro.)

F ➔ Not Administered Character: Ao definir um caractere para essa opção, o programa compreenderá que o respondente não chegou a realizar esse item. (Essa opção é muito utilizada quando se trata de Testagem Adaptativa Computadorizada [CAT], tendo em vista que os participantes raramente chegam às questões finais de um desses testes).

G ➔ Test for Differential Item Functioning: Essa característica também é opcional e deve ser marcada apenas se o usuário tiver interesse e informações que possam dividir o grupo em diferentes categorias para entender se há
diferenças significativas entre ambos (p.ex., homens e mulheres, brancos e pardos, ocidentais e orientais etc.). Como no exemplo, caso seja marcada, o programa irá gerar análises de DIF para homens (males) e mulheres (females).

3 – A aba *IRT Model*

Nessa aba, é possível que o usuário verifique qual análise irá realizar à luz da Teoria de Resposta ao Item. Veja a seguir as opções fornecidas por essa aba.

![Figura 3.1 – Opções da aba IRT Model](image)

**A ➪ Dichotomous Only:** Nessa opção, o teste que está sendo analisado apresentou apenas itens dicotômicos, ou seja, itens que podem ser corrigidos como certo e errado. Note que no exemplo apenas a coluna com modelos para itens dicotômicos é possível de ser marcada. Pode-se obter estatísticas para o modelo de 1, 2 e 3 parâmetros, que foram explicadas anteriormente.

**B ➪ Polytomous Only:** O pesquisador interessado em avaliar itens com mais de uma resposta possível, tal como a análise de um item avaliado através
da Escala Likert, deve marcar essa opção. Note que ao marcá-la a coluna da direita com os diversos modelos estará disponível para ser utilizada.

C → *Mixed Model (Dichotomous and Polytomous):* Quando o teste é formado por itens que podem ser avaliados como certo ou errado, e também por itens que podem ter mais de uma resposta possível, é necessário que o mesmo trabalhe com essa opção. Note que ao marcar essa opção será possível trabalhar com ambas as colunas. Ressalta-se a possibilidade de futuros trabalhos com itens politônicos, porém o presente se limita a trabalhar com itens dicotômicos. Nesse momento, será demonstrada a aba *Calibration.*

4 – A aba *Calibration*

Nessa aba, o pesquisador pode selecionar uma distribuição *a priori,* bem como opções de calibração usando os parâmetros da TRI. Veja a seguir quais as opções possíveis para serem seguidas nessa aba.

![Figura 4.1 – Características da aba Calibration](image)

**Figura 4.1 – Características da aba Calibration**

A → *Prior for the a parameter:* Essa opção permite o usuário fixar uma distribuição *a priori* específica para o parâmetro *a* (ou discriminação). Se o pesquisador for trabalhar com uma constante *D* igual a 1,0, deveria avaliar a
possibilidade de aumentar a média para a sua distribuição, tendo em vista que os valores desse parâmetro serão maiores para um fator de 1,7.

B ➔ Prior for the b parameter: Aqui o usuário pode definir média e desvio-padrão a priori para o parâmetro dificuldade do item.

C ➔ Prior for the c parameter: Nessa opção, estima-se distribuições de média e desvio-padrão a priori para o parâmetro acerto ao chute. Quando a média a priori variar para cada item de acordo com o número de alternativas, o pesquisador deverá selecionar a opção ao lado (use 1/#alt for the c prior mean).

As opções relativas à seção Calibration Options serão demonstradas abaixo.

D ➔ Perform no more than __ loops (E-M steps) during item parameter estimation: Essa opção é importante para ser utilizada quando o programa não encontrar convergência após realizar x loops, ou seja, após rodar o conjunto de dados algumas vezes, o programa deve encontrar padrões de repetição que caminham na mesma direção e sentido. Dessa forma, ele estará convergindo.

E ➔ Use a convergence criterion of __ for the item parameter estimates: Nessa área é possível estabelecer critérios de convergência para o programa parar de fazer interações. Há um número máximo que um loop possa ser processado que é computado como a soma dos valores absolutos de todos os parâmetros de itens; dessa forma, definir o critério para 0,001 irá fazer com que o programa gere 30 ou mais loops para alcançar esse critério de convergência.

F ➔ Use __ quadrature points during item parameter estimation: Permite ajustar o número de pontos de quadratura utilizados. Quanto mais pontos de quadratura se utiliza, mais precisas serão as estimativas, porém mais tempo o programa levará para rodar as análises.

G ➔ Item Parameter Scaling: Essa opção é necessária em apenas duas situações:

a) quando apenas itens dicotômicos estão sendo utilizados, e
b) usando o modelo de Rasch para itens dicotômicos.

Ao marcar a primeira opção, Center the dichotomous item parameters on theta, as estimativas de theta terão uma média zero e um desvio-padrão de 1. Ao clicar na opção Center the dichotomous item parameters on b, os valores de média e desvio-padrão para o parâmetro b serão de zero e um, respectivamente.

A seguir, são fornecidas informações acerca da aba Estimation.
5 – A aba *Estimation*

A figura 5.1 mostra as opções possíveis para obter as estatísticas da aba *Estimation*. Nessa aba, consegue-se observar quais são os métodos utilizados pelo programa para estabelecer o theta dos indivíduos. Explicações acerca do significado de cada um desses métodos fogem do escopo desse trabalho e podem ser encontradas no Manual apresentado no Apêndice A desta dissertação.

![Figura 5.1 – Informações fornecidas pela aba *Estimation*](image)

**A** → *Acceptable P(difficulty) range*: Nesse momento irão ser estimados os itens que poderão fazer parte da análise considerando o limite mínimo e máximo de sua dificuldade. Note que essa é uma estatística clássica e por isso apenas admite valores que possam ter sentido para a Teoria Clássica.

**B** → *Compute scale scores*: Ao marcar essa opção, o usuário fará com que as estimativas de theta para o teste inteiro e para cada domínio tenham uma média e desvio-padrão igual àqueles estabelecidos nas caixas. (No exemplo, o teste utilizado terá uma média de 50 e um desvio-padrão de 10.)
C ➔ Correct the item-total scores correlation for spuriousness: Ao deixar essa caixa marcada, o usuário estará corrigindo as correlações encontradas, tal que a correlação de um item com ele mesmo dentro de um teste não faça parte do resultado final. (Importante: Não é possível realizar essa correção para correlações entre o item e o theta, e deve-se admitir que spuriousness é um problema que testes com menos de 20 itens podem enfrentar, por diminuir muito a precisão das correlações encontradas).

D ➔ Exclude items from the calibration if the number of valid responses is less than __: Clicar nessa caixa fará com que os parâmetros da TRI não sejam calculados caso o item não apresente o critério mínimo de respostas para a realização de uma análise. Importante notar que itens omissos e não respondidos não fazem parte da análise.

A última aba a ser explorada no programa será a aba Output Options. Veja a seguir o que é possível realizar nela.

6 – A aba Output Options

A partir dessa aba é possível selecionar opções relativas ao modo com o qual o arquivo de saída será demonstrado, bem como diferenciar grupos com diferentes níveis de theta (p.ex., pessoas que alcançam e que não alcançam a nota desejada).
Figura 6.1 – Características da aba Output Options

A ➔ Perform two-group classification using a cutpoint of __ for the: Ao selecionar essa opção, o usuário irá estabelecer um ponto de corte que delimitará o grupo com theta alto e o grupo com theta baixo (p.ex.: se o ponto de corte for de 0,350, qualquer respondente que tenha alcançado esse valor de theta ficará no grupo com alta aptidão, enquanto aqueles que tiverem ficado abaixo desse ponto serão considerados de baixa aptidão).

B ➔ IRT Expected Proportion Correct (EPC): Essa opção também pode ser vista como o cálculo dos “escores verdadeiros” da TRI, e não estará disponível se existirem itens politômicos a serem avaliados. Em termos simples, é a proporção esperada de itens acertados por um indivíduo com uma aptidão estimada X.

C ➔ Theta Estimates: Pode ser colocada no grupo que não alcançou a aptidão necessária (low) e aquele que alcançou essa aptidão (high).

D ➔ Output Options: Todas as opções apresentadas nessa característica, quando marcadas, serão apresentadas no arquivo de saída por duas formas:

1) em um arquivo em que os dados são apresentados utilizando-se de vírgulas como separadoras de dados (CSV File), e

2) em um arquivo em que se utiliza de tabulações para separar os dados (Tab Delimited File).

Após uma apresentação breve das funções do programa, é necessário averiguar, por meio de um exemplo, como ele pode realizar algumas funções. Para a criação do exemplo, foi selecionada uma análise simples, utilizando o modelo Rasch ou modelo de 1 parâmetro. Serão demonstrados os passos para a obtenção dos resultados, desde a criação do arquivo de entrada até a visualização do arquivo de saída.

7- Exemplo

7.1 - Criando o arquivo de entrada
É necessário, como visto anteriormente, criar um arquivo de entrada para que seja possível realizar as análises desejadas. Veja a seguir o arquivo de entrada criado para o exemplo.

![Figura 7.1.1 – Arquivo de entrada de dados criado para o exemplo](image)

Note que nesse arquivo constam respostas para 12 itens para 10 respondentes. Como já citado acima ele é um arquivo delimitado por vírgulas e seus itens são dicotômicos (cujas questões podem estar certas ou erradas). Note que as abas apenas serão exploradas após a criação dos 3 arquivos necessários para as análises. Nesse momento é necessário que seja criado o arquivo de controle. (Importante: todos os arquivos são criados pelo bloco de notas e salvos no formato .txt, com exceção do arquivo de saída, que é criado pelo Microsoft Office Word e salvo no formato .rtf, que inicialmente deve estar em branco). Esse arquivo deverá ser inserido na lacuna Data Matrix file, da aba File.

7.2 - Criando o arquivo de controle

Relembrando, o arquivo de controle é o arquivo em que devem constar informações sobre os itens, tal como um “gabarito” do teste e informações que possam auxiliar o programa a julgar corretamente o que está sendo considerado no arquivo de entrada. Veja, a seguir, um arquivo de controle criado.

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>Y</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
<tr>
<td>Item</td>
<td>12</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>P</td>
</tr>
</tbody>
</table>
Figura 7.2.1 – Arquivo de controle estabelecido para os dados do exemplo.

Tendo em vista as explicações já fornecidas sobre esse arquivo, o programa irá interpretar que são 12 itens, que são avaliados dicotomicamente (por isso o número 1 na coluna logo após a identificação), com duas alternativas, avaliando o primeiro domínio, todos os itens serão incluídos na análise (mostrado pelo Y escrito para cada item) e todos os itens são numéricos começando em 0 (demonstrado pela letra P na última coluna). Esse arquivo deve ser inserido na lacuna Item control file da aba File. Para relembrar outras possibilidades de criação das características do arquivo de controle, consultar a seção 1.2.

7.3 – Criando o arquivo de saída de dados

A criação do arquivo de saída é a parte mais simples da análise. Para isso é necessário apenas que o pesquisador abra um arquivo do Word em branco, e utilize a opção “Salvar Como...” como mostrado na figura a seguir.

Figura 7.3.1- Utilizando o Word para salvar um arquivo de saída.

Logo após o usuário solicitar esse comando para salvar o arquivo em branco no local que desejar de seu computador, irá abrir uma caixa com opções para selecionar em que formato deseja salvar e deverá ser solicitado o formato Rich Text. Conforme explicado na figura a seguir.
Após a realização desses procedimentos, iremos iniciar as análises necessárias.

7.4 – Utilizando a aba **Files**

Nesse momento, será utilizada a aba **Files** para que sejam colocados os arquivos criados pelo usuário para que o programa consiga ler os mesmos. É necessário que os arquivos estejam corretamente construídos nesse momento para não ocorrer erros de leitura por parte do programa. Veja abaixo como e onde serão inseridos cada arquivo.

![Figura 7.4.1 – Arquivos de entrada, controle e saída inseridos no programa através da aba **Files.**](image)

Note que o arquivo de entrada criado irá ser inserido em **Data matrix file,** o arquivo de controle será inserido em **Item control file,** e o arquivo de saída será
inserido em *Output file*. Para inserir esses arquivos, o usuário deve clicar no botão e procurar o arquivo na pasta em que o mesmo foi salvo. Em *Run File*, está o nome do arquivo gerado com todas as informações necessárias.

A opção *Save the parameters in* foi marcada para que o programa salve os parâmetros dos itens em uma determinada extensão/formato. O formato escolhido foi o *ASC format (.par)*, que é a extensão utilizada com compatibilidade para programas de Testagem Adaptativa Computadorizada e, dessa forma, se o usuário quiser utilizar esse mesmo input para um Teste, não precisará salvar novamente.

A próxima etapa é marcar o que o programa fará na aba *Input Format*. Veja o que será necessário marcar para que seja possível rodar essa análise.

7.5 – Utilizando a aba *Input Format*

Essa aba, como visto anteriormente, fornece os limites para que o programa consiga estabelecer o que são dados relativos aos respondentes e quais dados são relativos aos itens em si. Veja a seguir os critérios que foram utilizados para o exemplo.

![Input Format](image)

**Figura 7.5.1 – Opções marcadas para a realização das análises do exemplo.**
O usuário pode notar que a opção selecionada *Delimited data* foi utilizada pois o arquivo construído possuía vírgulas que separavam as informações de respondentes e de itens.

**Delimited Data:**

- The data matrix file is delimited by a:
  - Comma
  - Tab
- Response matrix includes examinee ID in first column

Figura 7.5.2 – Opções *Delimited Data* marcadas para dados separados por vírgulas e cuja primeira coluna inclui informações do respondente.

A próxima etapa a ser realizada é a seleção do modelo de TRI que o usuário irá utilizar em suas análises. Veja na seção a seguir como selecionar esse modelo.

7.6 – Utilizando a aba *IRT Model*

Nessa aba, será selecionado o modelo com o qual se deseja trabalhar utilizando a TRI. Para fins de exemplo, iremos trabalhar com o modelo de um parâmetro (dificuldade) para itens dicotômicos. Veja a seguir o que será marcado nessa aba para que o programa realize essa análise.
Figura 7.6.1 – Opções utilizadas para a utilização do modelo de um parâmetro (dificuldade) para itens dicotômicos.

A imagem fornece ao usuário apenas a opção de trabalhar com o modelo de um parâmetro. É importante notar que a constante D é fixa para o modelo Rasch em um, por se tratar de um modelo logístico puro.

Ao selecionar o modelo desejado, o usuário pode calibrar seus parâmetros a priori. Utilizando a aba *Calibration*, o usuário é capaz de fixar distribuições que os parâmetros seguirão antes de ser iniciada qualquer análise. Veja a seguir as opções que serão marcadas nessa aba.

<table>
<thead>
<tr>
<th>Prior Distributions</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior for the a parameter</td>
<td>1.000</td>
<td>0.300</td>
</tr>
<tr>
<td>Prior for the b parameter</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Prior for the c parameter</td>
<td>0.250</td>
<td>0.030</td>
</tr>
</tbody>
</table>

**Figura 7.6.2 – Opções a serem trabalhadas em Prior Distributions**

As opções em *Prior Distributions* mostram ao usuário a partir de que média e desvio-padrão ele deseja que os dados estejam distribuídos para os 3 parâmetros. Como já mencionado anteriormente, é aconselhável aumentar a média do parâmetro a, caso se tenha $D = 1.0$. Porém, como estamos apenas trabalhando com o parâmetro $b$, deve-se levar em consideração apenas aquilo que está relacionado a esse parâmetro. Note que a média do parâmetro $b$ está em zero, ou seja, a dificuldade média do item é igual, inicialmente, para qualquer respondente. E o desvio-padrão é igual a 1.

Em *Calibration Options*, será mantido o que tiver sido marcado pelo programa, como é visto na figura a seguir.
Note que a opção *Center for dichotomous item parameters on b (Rasch model)* está marcada. Isso é um *Default* (padrão) do programa para rodar análises cuja principal característica a ser analisada é o parâmetro dificuldade do item. Para saber mais sobre as outra opções em *Calibration Options*, consultar a seção __.

O passo a seguir encontra-se na adequação da aba *Estimation* para a análise de dados.

7.7 – Utilizando a aba *Estimation*

A aba *Estimation* é utilizada para estabelecer critérios como dificuldade aceitável do item para ser inserido na calibração, correlação aceitável para ser incluído na análise, e qual método de estimação de theta será utilizado pelo programa. Por fugir do escopo do presente trabalho, não será explicado o que cada método de estimação pode fornecer ou como eles são feitos. Para maiores informações, consultar o Manual presente no Apêndice A da dissertação. Veja abaixo um exemplo da utilização da aba *Estimation*. 
As opções que surgem com o programa devem ser mantidas. Como explicado anteriormente, ao se descrever cada aba do programa, as opções apenas são modificadas se o usuário desejar incluir itens que podem vir a fugir do padrão encontrado pela maioria dos dados que seguem uma distribuição normal. Lembrando ainda que se o pesquisador estiver interessado em ter uma correlação mais “pura”, onde a correlação do item consigo mesmo no teste não influenciar as análises, deve marcar a caixa Correct the item-total score correlations for spuriousness.

A segunda parte da aba Estimation (Theta Estimation Method) é demonstrada a seguir.

O método mais utilizado para a estimação de theta com uma distribuição normal é o de máxima verossimilhança (Maximum Likelihood), dessa forma esse será também o método utilizado para o exemplo. Lembrando que se o usuário
estiver interessado em uma estimação para o teste completo e uma para cada domínio do teste deve marcar a caixa *Compute scaled scores*, isso fará com que o programa trabalhe com uma nova média e desvio-padrão estabelecidos pelo usuário.

A última aba *Output Format* analisada agora.

7.8 – Utilizando a aba *Output Format*

Essa aba tem por função auxiliar o usuário a selecionar o que quer que apareça no arquivo de saída e a forma com que a informação aparecerá. Veja a seguir as informações contidas nessa aba (*Classification* e *Output Options*).

![Image](image)

**Figura 7.8.1 – Parte Classification da aba Output Format**

Em *Classification*, é possível classificar dois grupos (passaram ou não passaram). Essa estatística apenas será solicitada se o usuário estiver trabalhando com um teste de caráter eliminatório. As informações nela contidas podem ser vistas na seção __ do presente trabalho. A parte *Output Options* é apresentada a seguir.
Figura 7.8.2 – Parte Output Options da aba Output Format.

Aqui, o usuário deve deixar marcado apenas aquilo que quiser que o programa forneça no arquivo de saída. E deve selecionar também se o arquivo foi delimitado por vírgula ou tabulação.

8 – UTILIZANDO O ARQUIVO DE SAÍDA

O usuário, após realizar as modificações necessárias, deverá rodar os arquivos. Para isso deve simplesmente clicar na opção que aparece no canto esquerdo inferior de qualquer aba. Caso não tenha erros nos comandos fornecidos ao programa, surgirá uma janela como a que se segue.

Figura 8.1 – Caixa de confirmação da finalização da análise e obtenção do arquivo de saída com os dados fornecidos pelo programa.
Note que ao clicar em “Yes”, o arquivo ainda não será aberto, mas a pasta destino onde o arquivo foi salvo será aberta. Dessa forma, o usuário deve apenas procurar o arquivo e dar um clique duplo sobre o mesmo para abri-lo.

9 – Interpretando os dados obtidos.

O arquivo de saída terá o formato de um relatório das análises solicitadas, porém com os padrões utilizados pela empresa criadora do software. Por isso, é importante que o usuário selecione apenas aquilo que lhe for de interesse e faça as correções de normatização para a publicação dos dados. As interpretações do arquivo de saída caberão ao usuário do teste e os escores e parâmetros podem ser definidos conforme o Manual do programa inserido no Apêndice B da presente dissertação.

10 – Conclusão

Este Manual integra a dissertação *Método para otimizar a aprendizagem no ensino médio usando o modelo de Rasch*. Ele é dirigido aos professores interessados em resultados mais precisos nas avaliações de seus alunos.

Ao empregar a TRI, especialmente o modelo de Rasch, é uma ferramenta que pode substituir a avaliação clássica, por sua maior precisão e detalhamento dos resultados.

Da mesma maneira ele pode orientar os especialistas curriculares, na seleção e ordenamento dos conteúdos, bem como a comunidade escolar participante dos projetos de avaliação nacional ENEM e SAEB, já inclusive adotantes da TRI.

Os dados podem ser processados pelo software Xcalibre, cujos procedimentos são aqui ilustrados passo a passo, para tornar mais fácil a tarefa.

José Maria F. de Farias

Dezembro de 2015
REFERÊNCIAS


Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. Psychometrika, 48, 233-246. Guyer, R. & Thomp