

## **Introdução à Ferramenta Teoria de Resposta ao Item – TRI**

Alexandra Waltrick Russi

Graduanda do curso de Estatística da Universidade Federal do Paraná.  
alexandra.w.russi@gmail.com

Setembrino Soares Ferreira Junior

Mestre em Administração. Professor do Departamento de Informática da Universidade Federal do Paraná.  
setembro@inf.ufpr.br

**Resumo:** A Teoria de Resposta ao Item (TRI) é, conforme Andrade (2000), definida como o conjunto de modelos matemáticos construídos para representar a probabilidade de um indivíduo dar certa resposta a um item em determinado teste. Para isto, leva em consideração parâmetros do item e a habilidade do respondente. Apesar dos primeiros estudos direcionados à TRI datarem da década de 60 do século passado, esta área de atuação ainda é pouco divulgada no país. O objetivo deste trabalho é apresentar os elementos básicos principais da TRI à comunidade Estatística, com o intuito de disseminar as possibilidades desta ferramenta. A elaboração deste instrumento de informação utilizou como base artigos, apostilas, apresentações de trabalhos em congressos e encontros, além da consulta à literatura clássica e alguns instrumentos de avaliação educacional. A pesquisa demonstrou que avaliações com a TRI apresentam forte ganho de qualidade e são um incentivo aos alunos para dedicarem mais esforços à multidisciplinaridade dos conhecimentos do que somente à memorização de “dicas” e “macetes”. Entender e desenvolver linhas de raciocínio significa preparar indivíduos do início do século XXI com mais qualidade e assertividade. Ao final, sugere-se temas para a elaboração de trabalhos futuros.

**Palavras-chave:** Instrumento de avaliação. Modelos de Resposta ao Item - MRI. Teoria de Resposta ao Item - TRI.

**Introdução:** A todo o momento, em várias situações e de várias formas, os conhecimentos dos indivíduos acerca de determinados assuntos são postos à prova. Alunos são classificados em disputas por uma vaga em Instituições de Ensino Superior (IES) a partir de seus desempenhos em avaliações. Candidatos concorrem a vagas de emprego na área pública e são classificados pelo número de acertos que atingem em provas dos concursos. Mas um ponto crucial que se deve levar em consideração na escolha ou na aprovação de um indivíduo nestes tipos de seleções é: Será que ele realmente tem o conhecimento que as questões aparentemente revelam? A avaliação é um importante instrumento de medição. Marcar uma opção em uma prova pode ser um ato de plena consciência do assunto em questão, ou apenas mais um item assinalado para não perder alguma chance de possível acerto. Para tentar chegar mais próximo de uma conclusão real acerca do verdadeiro conhecimento de um indivíduo foi desenvolvida, a partir da Teoria Clássica dos Testes (TCT), a Teoria de Resposta ao Item (TRI), definida como o conjunto de modelos matemáticos construídos para representar a probabilidade de um indivíduo dar certa resposta a um item em determinado teste. O objetivo deste trabalho é apresentar a TRI à comunidade de alunos de graduação em Estatística. Para isto, discorre-se sobre seus elementos básicos principais, sobre o modelo mais utilizado atualmente em testes de conhecimento, quais os parâmetros envolvidos, e o que o modelo pode revelar.

**Conteúdo:** Aplicadas à educação, psicologia, qualidade de vida ou à gestão da qualidade em indústrias e empresas, cada vez mais as avaliações vêm ocupando posição de destaque dentro das áreas de atuação das mais diversas profissões. Assim, a procura por melhoria da qualidade como um todo na execução dos processos de avaliação é uma realidade da atual sociedade do conhecimento. Avaliações têm como base de referência a Teoria Clássica dos Testes – TCT que, apesar de bastante antiga, ainda é muito utilizada. Na Psicometria, área da Psicologia que agrega as ciências exatas (notadamente Matemática Aplicada e Estatística) à Psicologia, a TCT é um instrumento para avaliação de testes, utilizada para mensurar um conjunto ou uma gama de comportamentos que se deseja conhecer melhor. Qualquer teste que se preste à validação e ao uso deve ser fruto de pesquisas psicométricas. Porém, a TCT apresenta algumas limitações, dentre as quais destacam-se: a) os parâmetros dependem da amostra na qual o instrumento foi aplicado; b) a discriminação (variáveis que melhor caracterizam uma amostra) é baseada no escore total do teste; c) o número de itens influencia na fidedignidade do instrumento; e d) há a necessidade de um número grande de participantes. Para melhorar a qualidade da ferramenta de avaliação, conhecimentos foram agregados à TCT e esta deu origem à Teoria de Resposta ao Item (TRI), em meados da década de 60 do século passado. Esta teoria considera que o comportamento humano é consequência de processos hipotéticos chamados de traços latentes. Estes podem ser entendidos como variáveis facilmente descritas e listadas (ex.: inteligência, ansiedade, habilidade na execução de um procedimento, nível de compreensão de um texto, etc.), porém não mensuráveis diretamente como o peso e a altura de uma pessoa. Para ilustrar e reforçar o conceito ora apresentado tomemos como exemplo uma avaliação de alunos do ensino fundamental: é razoável crer que alunos da 5ª série não terão bons escores se forem submetidos a provas do nível de conhecimento de alunos da 8ª série. Para acertar as questões desse teste, devem ter no mínimo conhecimentos ministrados na 8ª série. Essa quantidade de conhecimento, que pode ser por muitas vezes citada como habilidade, conforme ilustrado anteriormente, não é passível de mensuração direta. Mesmo que um aluno de 5ª série gabarite uma prova de 8ª série, é bem mais razoável assumir que a maioria de seus acertos tenham sido meros acasos. Essa habilidade, o tamanho do conhecimento acerca de um item, é o chamado Traço Latente. Com a finalidade de identificar o modelo mais adequado aos dados que se pretende avaliar, Andrade (2000) recomenda que seja determinada primeiramente a natureza do item que se deseja medir (dicotômico ou não-dicotômico). Itens com natureza não-dicotômica necessitam ser ajustados para dicotômicos, ademais da perda de significado de algumas respostas. Ainda segundo o mesmo autor, a seguir deve-se identificar o número de populações envolvidas (uma ou mais) e a quantidade de traços latentes a considerar (um ou mais). Verificada esta situação, ou seja, dados dicotômicos (originais ou ajustados), com apenas uma população e apenas um traço latente, nos remete à possibilidade de adoção de um dos modelos mais usuais da Teoria de Resposta ao Item – o Modelo Logístico unidimensional de 3 parâmetros (ML3). O ML3 leva em consideração: a) a dificuldade do item; b) a discriminação do item: diferença entre as médias do item, calculadas para os 27% que atingiram maiores escores no teste, dito grupo superior, e para os 27% com menores escores no teste, dito grupo inferior (SILVEIRA, 1982); e c) a probabilidade de uma resposta correta ter sido dada por indivíduos de baixa habilidade.

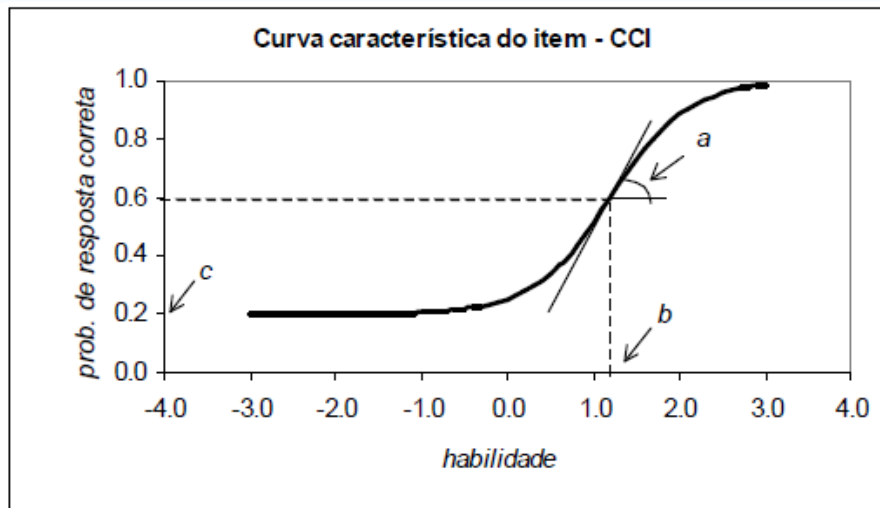
O índice de discriminação do item dá consistência (ou confiabilidade) ao teste.

A descrição do modelo ML3 é dada pela seguinte relação:

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \text{ com } i = 1, 2, \dots, k \text{ e } j = 1, 2, \dots, l$$

Onde:

$U_{ij}$  : Variável dicotômica que pode assumir: o valor 1 quando o indivíduo  $j$  responde corretamente o item  $i$ ; ou o valor 0 quando o indivíduo  $j$  não responde corretamente o item  $i$ ;  
 $\Theta_i$  : Traço latente do  $j$ -ésimo indivíduo;  
 $P(U_{ij} = 1 | \Theta_i)$  A probabilidade de um indivíduo  $j$ , com traço latente  $\Theta_i$ , responder corretamente o item  $i$ . Também chamada de Função de Resposta do Item – FRI.  
 $b_i$  : Parâmetro de dificuldade do item  $i$ , medido na mesma escala do traço latente.  
 $a_i$  : Parâmetro da distribuição do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item (CCI) no ponto  $b_i$ , ilustrada a seguir (Figura 1).



**Figura 1** – Representação da Curva Característica do Item.  
 Fonte: Andrade; Tavares; Valle (2000).

#### Curva Característica do Item (CCI)

É intuitivo que cada pessoa responda a um determinado item de acordo com o nível de suas habilidades implícitas. A probabilidade de um indivíduo acertar um determinado item dicotômico (com respostas do tipo “sim” ou “não”, por exemplo) será pequena se a habilidade da pessoa naquele assunto que o item se propõe a medir for pequena e vice-versa. Ao ser representada graficamente, tal probabilidade assume o formato de uma curva em forma de “S alongado”. A essa curva damos o nome de Curva Característica do Item (CCI), que é base na construção da TRI. Uma curva se diferencia de outra em função do nível de dificuldade e pelo poder de discriminação do item.

$c_i$ : Parâmetro do item que representa a probabilidade de indivíduos com baixo traço latente responderem corretamente um item. (Também referenciado como a probabilidade de acerto casual – o “chute”).

$D$ : Fator de escala: constante e igual a 1 para o caso de objetivar-se a semelhança entre os resultados gráficos obtidos no ML3 com os gerados por uma função ogiva normal. A função Ogiva é “abaixo de” ou crescente; assim, as frequências indicadas no gráfico são frequências acumuladas. Nas abscissas estão as classes e nas ordenadas as frequências.

Há ainda dois modelos primários que podem ser deduzidos a partir do ML3: o ML2 e o ML1. Historicamente, ambos ocorreram como passos anteriores ao ML3.

O Modelo Logístico unidimensional de 1 parâmetro, também conhecido como modelo de Rasch, leva em consideração somente o parâmetro da dificuldade do item, apresentado desta forma:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}, \text{ com } i = 1, 2, \dots, k \text{ e } j = 1, 2, \dots, l$$

Neste Modelo de Resposta ao Item (MRI), a possibilidade de acertar ao acaso é considerada  $c_i = 0$ . Além deste fato, todos os itens têm o mesmo poder de discriminação, ou seja, a dificuldade de um item relacionada ao traço latente do indivíduo não gera diferenciação entre itens (chances de acertos em questões de dificuldades distintas são iguais).

Caso seja considerada a diferença do poder de discriminação dos itens, teremos o Modelo Logístico unidimensional de 2 parâmetros, definido pela seguinte expressão:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \text{ com } i = 1, 2, \dots, k \text{ e } j = 1, 2, \dots, l$$

O Ministério da Educação (MEC), órgão do Poder Executivo brasileiro responsável pelas políticas e diretrizes da educação no país, para melhorar a qualidade de suas avaliações de ensino, vem utilizando a TRI para o processamento de suas provas e geração dos escores dos alunos. Um caso típico é o do processamento das notas do ENEM, onde os escores (resultados obtidos) são bem mais que a simples contagem das questões marcadas como certas. O MEC usa a TRI sob o Modelo Logístico de 3 parâmetros pelo fato deste permitir a comparação de resultados entre diversos ciclos de avaliação. Com esta nova modalidade de avaliação de escores, o aluno é incentivado, em tese, a ater-se mais à multidisciplinaridade dos conhecimentos do que somente a memorizar “dicas” e “macetes”, seguindo o estilo arcaico de preparação para provas e concursos. Neste processo, entender e desenvolver linhas de raciocínio dá mais garantias de acerto das questões. É claro que o “chute” não deixará de ocorrer, mas será substituído por uma opção mais coerente de resposta de acordo com a interpretação particular do indivíduo respondente. Segundo Torrecilla (2006), os fatores que realmente garantem o sucesso do teste, propiciados pela utilização da TRI, são: a) a elaboração de provas constituídas por itens independentes dos grupos de aplicação; b) a garantia de que a habilidade do aluno seja traduzida por um escore independentemente da dificuldade do teste; c) a construção de itens relacionados perfeitamente à habilidade que se pretende medir; e d) a garantia de confiabilidade e comparabilidade dos resultados numa perspectiva espacial e temporal, além de não se fazer necessário o uso de provas estritamente paralelas.

Para a utilização do ML3 e a estimação de seus parâmetros é necessário recorrer-se ao apoio de programas computacionais estatísticos. A literatura consultada enumerou, entre outros, os seguintes: BILOG, Action+R, BIMAIN, MICROCAT, TESFACT e MICROFACT. O conteúdo e a operação destes programas fogem ao escopo deste trabalho. Cumpre citar, no entanto, que os programas se propõem a estimar, analisar e gerar dados capazes de explicar (ou próximo disto) a população respondente, na forma de gráficos e relatórios para análise.

**Conclusões:** Desenvolver avaliações com qualidade compatível com a demanda exigida pelas diversas áreas de atuação de empresas e profissionais da sociedade do início do século XXI é mais um desafio que se apresenta à comunidade de Estatística. Este requer a possibilidade de fornecer-se uma melhor avaliação de escores em provas e testes, indo além da simples visão de que o escore deve ser baseado somente nas questões assinaladas. Este trabalho discorreu sobre uma possível resposta àquela demanda através da apresentação dos conceitos básicos principais da Teoria de Resposta ao Item (TRI). Através dele, esta teoria poderá vir a ser mais

conhecida no meio acadêmico. A elaboração deste instrumento de informação utilizou como base artigos, dissertações de mestrado, apostilas, apresentações de trabalhos em congressos e encontros, além da consulta à literatura clássica. Alguns instrumentos de avaliação educacional que utilizaram a TRI, como o Sistema de Avaliação da Educação Básica (SAEB), o Exame Nacional de Desempenho dos Estudantes (ENADE) e o Exame Nacional do Ensino Médio (ENEM) também foram consultados. A TRI já é uma realidade nas instituições responsáveis por estas avaliações. À medida que o contato do meio acadêmico com esta teoria seja antecipado ocorrerá a mais rápida disseminação de seu uso, o que pode trazer uma inestimável contribuição à sociedade, por tratar-se de um elemento diferencial na busca do homem por conhecer melhor a realidade que o cerca, através de avaliações mais criteriosas. O presente trabalho não abordou de forma exaustiva a TRI, pela amplitude do tema e pela restrição de espaço físico. Muito há para se desenvolver: exploração das especificações, funcionalidades e limitações dos programas computacionais estatísticos de apoio ao uso da TRI; aplicações educacionais de apoio a professores para a estimação da aprendizagem de conteúdos e ensaios de pequeno porte em situações controladas (exemplo: melhoria na abordagem de assuntos de uma disciplina de um curso de graduação) são trabalhos que certamente contribuirão para o ganho de conhecimento dos profissionais de Estatística e para o melhor desempenho dos profissionais de outras áreas no desenrolar de suas funções.

#### Referências:

ALEXANDRE, J. W. C.; ANDRADE, D. F.; VASCONCELOS, A. P.; ARAUJO, A. M. S. Uma proposta de análise de um construto para a medição dos fatores críticos da gestão pela qualidade por intermédio da teoria de resposta ao item. **Gestão da Produção**, v. 9, n. 2. São Carlos: UFSCAR, 2002.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria da resposta ao item: conceitos e aplicações**. Apostila, Associação Brasileira de Estatística – ABE, 2000

Assessoria de Imprensa do Inep / MEC. **Para entender a nota do ENEM**. Sítio eletrônico do Ministério de Educação. Disponível em: <[http://www.inep.gov.br/imprensa/noticias/enem/news10\\_04.htm](http://www.inep.gov.br/imprensa/noticias/enem/news10_04.htm)>. Acessado em 29 de março de 2010.

AZEVEDO, C. L. N.; GEMERMAN, D. **Introdução à teoria de resposta ao item: Minicurso introdutório**. In: 1º Congresso Brasileiro de Teoria de Resposta ao Item (COMBRATRI), 65. 2009. **Anais...** Florianópolis, 9 a 11 de dezembro de 2009.

BORTOLOTTI, S. L. V. **Aplicação de um modelo de desdobramento graduado generalizado da teoria de resposta ao item – tri**. 2003. Dissertação (Mestrado em Engenharia de Produção com ênfase em Estatística). Universidade Federal de Santa Catarina, Florianópolis, 2003.

CARVALHO, L. F. **Primeiros passos... Teoria de Resposta ao Item**. Vídeo. Disponível em: <[http://www.youtube.com/watch?v=E3eI9j\\_MhBA](http://www.youtube.com/watch?v=E3eI9j_MhBA)>. Acessado em 10 de abril de 2010.

CASAL, G. B. **Manual de evaluación psicológica: fundamentos, técnicas y aplicaciones**. 1.

ed. Madrid: Siglo XXI, 1997.

PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item – TRI. **Avaliação Psicológica**, v. 2, n. 2. p. 99-110, 2003.

PITON, J. **Teoria de Resposta ao Item e o novo ENEM**: esclarecendo alguns pontos para estudantes. Blog. Disponível em: <[http://jpiton.blogspot.com/2009/05/teoria-de-resposta-ao-item-e-o-novo\\_9153.html](http://jpiton.blogspot.com/2009/05/teoria-de-resposta-ao-item-e-o-novo_9153.html)>. Acessado em 28 de março de 2010.

SILVEIRA, F. L. Considerações sobre o índice de discriminação de itens em testes psicrométricos e educacionais. **Educação** (Porto Alegre, RS, PPGE, PUCRS), n. 5, p.31-37, 1982.

TORRECILLA, F. J. M. **Estudios sobre eficacia escolar en Iberoamérica**: 15 buenas investigaciones. 1. ed. Bogotá: Convenio Andrés Bello, 2006.