

**UNIVERSIDADE FEDERAL DO AMAZONAS
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
DEPARTAMENTO DE APOIO À PESQUISA
PROGRAMA INSTITUCIONAL DE INICIAÇÃO CIENTÍFICA**

**ENTENDENDO A TEORIA DE RESPOSTA AO ITEM APLICADA AO
ENEM**

Bolsista: Gleici Jane Sena Cruz, CNPq

MANAUS

2010

**UNIVERSIDADE FEDERAL DO AMAZONAS
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
DEPARTAMENTO DE APOIO À PESQUISA
PROGRAMA INSTITUCIONAL DE INICIAÇÃO CIENTÍFICA**

**RELATÓRIO FINAL
PIB-E-0028/2010
ENTENDENDO A TEORIA DE RESPOSTA AO ITEM APLICADA AO
ENEM**

**Bolsista: Gleici Jane Sena Cruz, CNPq
Orientadora: Prof^a. Dr^a. Maria Ivanilde Silva Araujo**

**MANAUS
2011**

RESUMO

Todos os anos milhões de pessoas se submetem a exames de seleção com o intuito de ingressar em uma instituição de ensino superior no Brasil e essa disputa se dá principalmente por vagas em universidades federais. O ENEM, em 2009, tornou-se a principal forma de ingresso nas instituições federais de ensino superior através do processo seletivo unificado. Fazendo-se necessário desenvolver avaliações com qualidade houve mudanças na metodologia de avaliação, agora não mais baseada somente no escore das questões assinaladas. Para tal o MEC substituiu a Teoria Clássica pela TRI, metodologia que dá suporte à elaboração e correção das provas e é utilizada em vários países no campo de teste.

O presente trabalho visa apresentar a TRI a comunidade acadêmica da UFAM, os elementos básicos principais, os modelos mais utilizados e parâmetros, fazer uma análise dos diferentes modelos da TRI e suas diferentes parametrizações e identificar como a Teoria de Resposta ao Item avalia os alunos do ENEM, processo atual de ingresso de 50% dos alunos da UFAM

Os primeiros modelos da TRI datam da década de quarenta, e está é definida segundo ANDRADE (2000) como um conjunto de modelos matemáticos construídos para representar a probabilidade de um indivíduo dar certa resposta a um item como função dos parâmetros do item e do traço latente do respondente na área do conhecimento avaliada. Quanto maior a habilidade do respondente, maior a probabilidade de acerto. Os vários modelos propostos na literatura para estimar os parâmetros da TRI dependem fundamentalmente de três fatores: a natureza do item, o número de populações envolvidas e a quantidade de traços latentes que está sendo medida. O Modelo de TRI adotado pelo ENEM é o Modelo Logístico de três parâmetros para dados dicotômicos. Esta metodologia vem sendo amplamente empregada em avaliações educacionais no Brasil como o Sistema de Avaliação da Educação Básica (SAEBE) e o ENEM por poder comparar diferentes grupos numa mesma escala de habilidades - processo chamado equalização e ser uma forma de avaliação mais criteriosa.

SUMÁRIO

RESUMO	3
1. INTRODUÇÃO	5
2. REVISÃO BIBLIOGRÁFICA.....	6
3. METODOLOGIA	7
4. RESULTADOS E DISCURSSÕES.....	8
4.1. Modelos para itens dicotômicos envolvendo um único grupo	8
4.2. Modelos para itens não dicotômicos envolvendo um único grupo	9
4.3. Estimação	10
4.4. Equalização.....	12
4.5. Escala de Habilidade	12
4.6. Recursos Computacionais	12
5. CONCLUSÕES.....	13
REFERÊNCIAS	14

1. INTRODUÇÃO

O Exame Nacional do Ensino Médio (ENEM), criado em 1998 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), do Ministério da Educação, é um exame individual e de caráter voluntário, oferecido anualmente aos concluintes e egressos do ensino médio, com o objetivo principal de possibilitar uma referência para autoavaliação, a partir das competências e habilidades que o estruturam e como forma de acesso ao ensino superior (FERREIRA, 2009).

Todos os anos milhões de pessoas se submetem a exames de seleção com o intuito de ingressar em uma instituição de ensino superior no Brasil e essa disputa se dá principalmente por vagas em universidades federais. Visando unificar a seleção às vagas das Instituições Federais de Ensino Superior (IFES) por meio de uma única prova, possibilitar a mobilidade acadêmica e reestruturar os currículos do ensino médio, o Ministério da Educação, em 2009, reformulou o Exame Nacional do Ensino Médio (ENEM). Esta reformulação substituindo a Teoria Clássica (TC), metodologia de avaliação utilizada até então, onde os resultados obtidos se baseiam em escores brutos ou padronizados, pela a Teoria de Resposta ao Item (TRI), técnica muito empregada em vários países no campo de testes.

A TRI, segundo ANDRADE (2000) propõe modelos para as características do indivíduo que não podem ser observadas diretamente, chamadas traços latentes. É definido como um conjunto de modelos matemáticos construídos para representar a probabilidade de um indivíduo dar certa resposta a um item como função dos parâmetros do item e do traço latente do respondente na área do conhecimento avaliada. Quanto maior a habilidade do respondente, maior a probabilidade de acerto. A partir de respostas dadas por indivíduos a um conjunto de itens deseja-se estimar os parâmetros dos itens (calibração), a habilidade ou proficiência do aluno e a proficiência média de um grupo de alunos. Os parâmetros dos itens e as proficiências dos indivíduos são invariantes, exceto pela escolha de origem e escala

O presente trabalho visa apresentar a TRI a comunidade acadêmica da UFAM, os elementos básicos principais, os modelos mais utilizados e parâmetros, fazer uma análise dos diferentes modelos da TRI e suas diferentes parametrizações e identificar como a Teoria de Resposta ao Item avalia os alunos do ENEM, processo atual de ingresso de 50% dos alunos da UFAM, através do sistema seleção unificada, cujo conhecimento não é bem assimilado na comunidade acadêmica e seu estudo está fora do atual currículo do curso de Estatística da UFAM.

2. REVISÃO BIBLIOGRÁFICA

Os primeiros modelos da TRI datam da década de quarenta com o trabalho de Lawley, que estimou os parâmetros da curva característica do item (CCI) a partir das estimativas de máxima verossimilhança (COSTA, 2001). Inicialmente estes modelos foram desenvolvidos na forma de uma função ogiva normal, os testes eram corrigidos de forma dicotômica e apenas uma habilidade de um só grupo era medida.

O desenvolvimento do modelo unidimensional de dois parâmetros baseado na função ogiva normal se deu na década de cinquenta com as publicações de Frederic Lord (1952), que posteriormente sentindo a necessidade de incorporar um parâmetro que tratasse do problema do acerto casual desenvolveu o modelo unidimensional de três parâmetros (LORD, 1952 apud ARAUJO, ANDRADE & BORTOLOTTI, 2009, p. 4). Em 1968 Birnbaum substituiu em ambos os modelos a função ogiva normal pela função logística, matemática mais conveniente, pois é uma função explícita dos parâmetros dos itens e não depende de integração (ANDRADE, 2000). Rasch (1960) propôs paralelamente ao trabalho de Lord o modelo unidimensional de um parâmetro, expresso também na da função ogiva normal, sendo este modelo mais tarde descrito por Wright por um parâmetro logístico. (RASH, 1960 apud ARAUJO, ANDRADE & BORTOLOTTI, 2009, p. 4).

Em 1969 com objetivo de obter maiores informações acerca das respostas dos indivíduos e não apenas se responderam corretamente ou incorretamente os itens, Samegima propôs o Modelo de Resposta Gradual. Bock (1972) desenvolveu o Modelo de Resposta Nominal e Masters (1982) propôs o Modelo de Crédito Parcial. Outros modelos para respostas que não fossem consideradas exclusivamente como dicotômicas, com diferentes estruturas entre essas categorias foram propostos por Andrich (1978), Muraki (1992), entre outros. (SAMEGIMA, 1969; BOCK, 1972; MASTERS, 1982; ANDRICH, 1978; MURAKI, 1992 apud ARAUJO, ANDRADE & BORTOLOTTI, 2009, p. 4).

Os modelos logísticos de 1, 2 e 3 parâmetros para mais de uma população respondente foram introduzidos recentemente por Bock & Zimowskil (1997), o que possibilitou a comparação de rendimento de duas ou mais populações submetidas a diferentes testes com itens comuns (BOCK & ZIMOWSKI, 1997 apud ARAUJO, ANDRADE & BORTOLOTTI, 2009, p. 4). Estes modelos não serão abordados neste trabalho. No Brasil, a TRI foi usada pela primeira vez em 1995 na análise dos dados do Sistema Nacional de Ensino Básico – SAEB. (ANDRADE & KLEIN, 1999 apud ARAUJO, ANDRADE & BORTOLOTTI, 2009, p. 4).

3. METODOLOGIA

Os vários modelos propostos na literatura para estimar os parâmetros da TRI dependem fundamentalmente de três fatores:

- 1- da natureza do item – de forma dicotômica ou não dicotômica;
- 2- do número de populações envolvidas – apenas uma ou mais de uma;
- 3- da quantidade de traços latentes que está sendo medida – um ou mais.

O processo de amostragem é aleatório simples e nesse trabalho estaremos considerando apenas os modelos unidimensionais, onde apenas uma habilidade ou traço latente está sendo medida.

Os modelos da TRI se diferenciam na forma matemática da função característica do item e/ou no número de parâmetros envolvidos na modelo e dependem da natureza do item e do processo de resposta.

Nos modelos para itens dicotômicos destacam-se os modelos logísticos de 1, 2 e 3 parâmetros por serem os mais utilizados. Será dada ênfase apenas ao modelo logístico de três parâmetros por ser o modelo empregado no ENEM e o mais completo, sendo que os outros dois modelos podem ser obtidos a partir deste.

Para os itens não dicotômicos não se considera somente se o indivíduo respondeu corretamente ou não o item, mas também a resposta dada por ele. Temos os seguintes modelos: Modelo de Resposta Nominal, Modelo de Resposta Gradual, Modelo de Escala Gradual, Modelo de Crédito Parcial e Modelo de Crédito Parcial Generalizado. Será abordado apenas os modelos Nominal e de resposta Gradual, pois os demais são variantes destes dois modelos.

De acordo com a Assessoria de Comunicação Social do Ministério da Educação (2010) a nova prova do ENEM será estruturada a partir de uma matriz de habilidades e um conjunto de conteúdos associados a elas.

O novo exame está dividido em quatro áreas de conhecimento, sendo um teste para cada área. Estas são: Linguagens, Códigos e suas Tecnologias (incluindo redação); Ciências Humanas e suas Tecnologias; Ciências da Natureza e suas Tecnologias e Matemática e suas Tecnologias, permitindo ao Enem ter boa precisão na avaliação dos traços latentes. Em relação à escala será construída quatro escalas distintas, uma para cada área do conhecimento capaz de ordenar os estudantes conforme seu nível de proficiência, sendo possível às IFES estabelecer distintas ponderações ou pontos de corte para seleção de seus candidatos.

4. RESULTADOS E DISCURSÕES

4.1. Modelos para itens dicotômicos envolvendo um único grupo

Para os itens dicotômicos os modelos logísticos de 1, 2 e 3 parâmetros são os mais utilizados e se diferenciam pelo número de parâmetros que utilizam para descrever o item. O modelo logístico de um parâmetro considera apenas a dificuldade do item, já o modelo logístico de dois parâmetros considera a dificuldade do item e a discriminação, o modelo logístico de três parâmetros considera além da dificuldade do item e da discriminação a probabilidade do acerto casual.

O modelo Logístico de três parâmetros (ML3) é o modelo mais utilizado pela TRI, sua forma matemática é dada por:

$$P(U_{ij} = 1/\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-D\alpha_i(\theta_j - b_i)}}$$

onde:

$i = 1, 2, \dots, I$ representa os I itens propostos para avaliar o traço latente observado.

$j = 1, 2, \dots, n$ representa os n indivíduos que compõe a amostra.

U_{ij} é uma variável dicotômica que assume os valores 1 se o indivíduo j responder corretamente o item e 0 caso contrário.

θ_j representa o traço latente ou habilidade do j -ésimo indivíduo.

$P(U_{ij} = 1/\theta_j)$ é a Função de Resposta do Item e representa a probabilidade do indivíduo j com habilidade θ_j responder corretamente o item i .

b_i é o parâmetro de discriminação do item i , medido na mesma escala da habilidade.

α_i é o parâmetro de discriminação do item i , com valor proporcional à inclinação da Curva Característica do Item no ponto b_i .

c_i é o parâmetro do item que representa a probabilidade de acerto casual.

D é um fator de escala, constante e igual a 1. Assume valor 1,7 quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal.

A relação existente entre $P(U_{ij} = 1/\theta_j)$ e os parâmetros do modelo é mostrada pela Curva Característica do Item (CCI), representada na Figura 1, sendo que o modelo baseia-se no fato de que indivíduos com maior habilidade possuem maior probabilidade de acertar o item, sendo esta relação não linear.

Podemos observar na Figura 1, que o gráfico tem forma de "s" com inclinação e deslocamento na escala de habilidade definidos pelos parâmetros do item, sendo a escala de habilidade arbitrária.

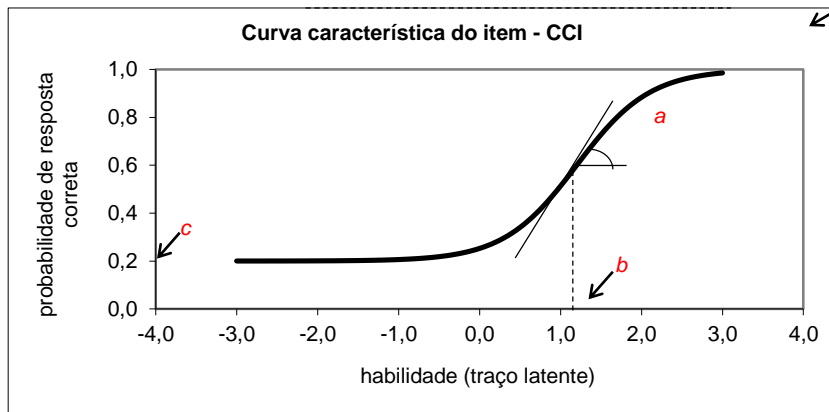


Figura 1: Curva Característica do Item

Segundo SAMEJIMA (1969), a CCI descreve a relação entre a mudança do traço latente com a mudança na probabilidade de uma dada resposta sendo a CCI um dos pressupostos relevantes nos modelos da TRI. O outro pressuposto é a independência local, ou seja, deve haver independência entre as respostas dadas aos diferentes itens do teste para uma dada habilidade e este conjunto de itens deve ser homogêneo.

4.2. Modelos para itens não dicotômicos envolvendo um único grupo

Os modelos para os itens não dicotômicos são empregados para análise de itens abertos e itens de múltipla escolha que não são avaliados de forma graduada.

O Modelo de Resposta Nominal foi desenvolvido por Bock em 1972 com base no modelo logístico de dois parâmetros para avaliar itens de múltipla escolha considerando todas as categorias de resposta. Sua função matemática é dada por:

$$P(U_{ijs} = 1 | \theta_j) = \frac{\exp[a_{is}(\theta_j - b_{is})]}{\sum_{h=1}^{m_i} \exp[a_{ih}(\theta_j - b_{ih})]}$$

Com $i = 1, 2, \dots, l$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m_i$ onde m_i são as opções avaliáveis e a_{is} e b_{is} iguais como no modelo logístico estando relacionados com a k -ésima opção.

O Modelo de Resposta Gradual foi desenvolvido por Samejima em 1969 e assume que as categorias de resposta de um item podem ser ordenadas entre si, busca também obter mais informações acerca das respostas dos indivíduos e sua função matemática é dada por:

$$P(U_{ijs} = 1 | \theta_j) = \frac{1}{1 + \exp[-a_i(\theta_j - b_{is})]} - \frac{1}{1 + \exp[-a_i(\theta_j - b_{i(s+1)})]}$$

Com $b_{i1} \leq b_{i2} \leq \dots \leq b_{im_i}$

A ordenação dos níveis de dificuldades das categorias de um item é dada de acordo com a classificação de seus escores e é denotado por $k = 0, 1, \dots, m_i$ onde $(m_i + 1)$ é o número de categorias do i -ésimo item, onde o parâmetro de inclinação do item e os m_i

valores de dificuldades precisam ser estimados. Logo para cada item, o número de parâmetros a ser estimado será dado pelo número de categorias de respostas.

4.3. Estimação

Uma das etapas da TRI é a estimação dos itens e das habilidades dos respondentes, processo conhecido como calibração, onde geralmente apenas as respostas dos indivíduos aos itens são conhecidas. A calibração pode ser realizada de três formas: estimação das habilidades, quando os parâmetros dos itens já são calibrados; estimação dos parâmetros dos itens, quando já se conhece as habilidades e a situação mais comum, estimação dos parâmetros dos itens e das habilidades dos indivíduos simultaneamente. Para a última situação há duas abordagens usuais: estimação conjunta ou em duas etapas, primeiro a estimação dos parâmetros dos itens e depois a estimação das habilidades.

A estimação é feita pelo Método de Máxima Verossimilhança, cujo objetivo é encontrar o conjunto de valores para os parâmetros que maximizem a verossimilhança das respostas dadas aos itens, através de um processo iterativo, como o algoritmo de Newton-Raphson, Scoring de Fisher ou alguns procedimentos bayesianos. Nesses métodos algumas notações e suposições fazem-se necessárias. $U_j = (U_{j1}, U_{j2}, \dots, U_{ji})$ o vetor de respostas do indivíduo j , $U_{..} = (U_{1.}, U_{2.}, \dots, U_{n.})$ o conjunto de respostas integral e as observações serão representadas por u_{ji}, u_j e u ; $\theta = (\theta_1, \dots, \theta_n)$ o vetor de habilidades dos n indivíduos, $\zeta = (\zeta_1, \dots, \zeta_n)$ o conjunto de parâmetros dos itens a_i, b_i e c_i dos n indivíduos e as seguintes suposições: as respostas de indivíduos diferentes são independentes e independência local, ou seja, para uma dada habilidade as respostas aos itens são respondidas de forma independente por cada indivíduo.

Usando as suposições, podemos escrever a verossimilhança, $L(\zeta) = P(U = u_{..}, \zeta)$, para a estimação dos parâmetros dos itens dado que as habilidades são conhecidas para um ML3 como

$$L(\zeta) = \prod_{j=1}^n P(U_j = u_j | \theta_j, \zeta) = \prod_{j=1}^n \prod_{i=1}^I P(U_{ji} = u_{ji} | \theta_j, \zeta_i)$$

onde na ultima igualdade usamos o fato de que a distribuição de U_{ij} só depende de ζ através de ζ_i . Usando a notação $P_{ji} = P(U_{ji} = u_{ji} | \theta_j, \zeta_i)$ e $Q_{ji} = 1 - P_{ji}$, temos que

$$P(U_{ji} = u_{ji} | \theta_j, \zeta_i) = P(U_{ji} = u_{ji} | \theta_j, \zeta_i)^{u_{ji}} P(U_{ji} = u_{ji} | \theta_j, \zeta_i)^{1-u_{ji}} = P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}$$

Substituindo

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}$$

A log-verossimilhança será dada por:

$$\log L(\zeta) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}.$$

Os Estimadores de Máxima Verossimilhança (EMV) de ζ_i são obtidos a partir da derivada da log-verossimilhança igualada à zero.

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) - (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right) \right\} = \sum_{j=1}^n \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

Por conveniência, consideremos a seguinte ponderação: $W_{ji} \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}$, onde:

$$P_{ji}^* = \left\{ 1 + e^{-D a_i (\theta_j - b_i)} \right\}^{-1} \text{ e } Q_{ij}^* = 1 - P_{ij}^*$$

Reescrevendo a equação acima temos:

$$\frac{\partial \log L(\zeta)}{\partial \zeta_i} = \sum_{j=1}^n \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \zeta_i} \right)$$

Onde

$$\frac{\partial P_{ji}}{\partial a_i} = D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^*; \quad \frac{\partial P_{ji}}{\partial b_i} = -D a_i (1 - c_i) P_{ji}^* Q_{ji}^* \text{ e } \frac{\partial P_{ji}}{\partial c_i} = Q_{ji}^*$$

Os Estimadores de Máxima Verossimilhança para os parâmetros a_i , b_i e c_i são, dados respectivamente pelas equações

$$a_i: D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji})(\theta_j - b_i) W_{ji} = 0; \quad b_i: -D a_i (1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} = 0;$$

$$c_i: \sum_{j=1}^n (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} = 0.$$

De forma análoga estimamos as habilidades dado que os parâmetros dos itens são conhecidos, agora escrevemos a verossimilhança como função de θ e não de ζ . O Estimador de Máxima Verossimilhança será obtido a partir de

$$\frac{\partial \log L(\theta)}{\partial \theta_j} = \sum_{i=1}^I \left\{ (u_{ij} - P_{ij}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_i} \right)$$

Onde

$$\frac{\partial P_{ji}}{\partial \theta_i} = D a_i (1 - c_i) P_{ji}^* Q_{ji}^*$$

O Estimador de Máxima Verossimilhança para o parâmetro θ_j é dado pela equação

$$\theta_j: D \sum_{i=1}^n a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ij} = 0.$$

Estas equações não possuem solução explícitas e por isso faz-se necessário algum método iterativo para a obtenção das estimativas de máxima verossimilhança dos parâmetros dos itens e das habilidades dos respondentes. Uma das formas de estimação

conjunta dos parâmetros dos itens pode ser feita individualmente através de um método iterativo usando estimativas iniciais para as habilidades θ (escores padronizados, por exemplo) e tratando essas habilidades como conhecidas, estima-se ζ_i , com estimativas iniciais para ζ e tratando estes parâmetros como conhecidos, estimamos as habilidades θ_j . Este procedimento é realizado em dois estágios e é repetido até a convergência das habilidades e dos parâmetros dos itens.

A principal diferença da estimação conjunta se dá no tratamento da escala em que todos os parâmetros são identificados. Na estimação conjunta não há uma escala definida, e portando, devemos estabelecê-la devido a não-identificabilidade do modelo, onde mais de um conjunto de parâmetros produz o mesmo valor no ML3, e conseqüentemente na verossimilhança. Essa não-identificabilidade pode ser eliminada especificando-se uma medida de posição e outra de dispersão para as habilidades, definindo assim uma escala para as habilidades e conseqüentemente para os parâmetros dos itens.

4.4. Equalização

KOLEN e BREANN (1995) definem “equalizar” como equiparar, tornar equiparável, o que na TRI significa colocar parâmetros de itens vindo de pesquisas distintas ou habilidades de respondentes de diferentes grupos numa escala comum, tornando os itens e/ou as habilidades comparáveis. A TRI não pode ser usada quando grupos fazem diferentes provas totalmente distintas. O conjunto de itens a ser estimados pode ser composto de itens novos (itens que ainda não foram calibrados), itens já calibrados ou itens novos e calibrados.

4.5. Escala de Habilidade

As escalas de habilidades tornam possível a interpretação pedagógica dos valores das habilidades. São definidas por *níveis âncora*, que são pontos selecionados pelo analista na escala de habilidades para serem interpretados pedagogicamente, estes por sua vez são caracterizados por conjuntos de itens denominados *itens âncora*, itens selecionados bastante acertados por indivíduos com aquele nível de habilidade e pouco acertados por indivíduos um nível de habilidade imediatamente inferior para cada um dos *níveis âncora*.

4.6. Recursos Computacionais

A literatura consultada enumerou os seguintes programas computacionais: TESTFACT, PARSCALE, BILOG e BILOG-MG usados atualmente no Brasil para a utilização do ML3 e a estimação de seus parâmetros abordados pela TRI. O conteúdo e a operação destes programas fogem ao escopo deste trabalho.

5. CONCLUSÕES

O ENEM, em 2009, tornou-se a principal forma de ingresso nas instituições federais de ensino superior através do processo seletivo unificado. Fazendo-se necessário desenvolver avaliações com qualidade houve mudanças na metodologia de avaliação, agora não mais baseada somente no escore das questões assinaladas. Para tal o MEC substituiu a Teoria Clássica pela TRI, metodologia que dá suporte à elaboração e correção das provas e é utilizada em vários países no campo de teste.

Segundo o sítio “Ultimo Segundo” (2011), os especialistas enumeram vantagens na sua utilização para elaboração de exames como o Enem: provas aplicadas em momentos diferentes podem ser comparadas, o conhecimento de cada aluno é avaliado de forma mais precisa e não há necessidade de aplicar as provas no mesmo dia para milhares de candidatos.

O Modelo de TRI adotado pelo ENEM é o Modelo Logístico de três parâmetros, usado para descrever o item – que é o foco da teoria, onde são consideradas a dificuldade do item, a discriminação e a probabilidade de acerto casual. Estes parâmetros são estimados por máxima verossimilhança, mas como as equações não possuem forma analítica são utilizados métodos iterativos como o Scoring de Fisher. São empregados também métodos bayesianos para a estimação dos parâmetros.

Os diferentes níveis de conhecimento são medidos através dos itens, estes variam de fáceis, medianos e difíceis, e separam quem sabe o conteúdo de quem acertou ao acaso. Estimado os parâmetros dos itens, define-se a escala de habilidade para as interpretações pedagógicas e descreve-se a relação entre a mudança da habilidade com a mudança na probabilidade de uma dada resposta através da Curva Característica do Item. Sendo que no ENEM há quatro escalas de habilidade, uma para cada área do conhecimento.

Esta metodologia vem sendo amplamente empregada em avaliações educacionais no Brasil como o Sistema de Avaliação da Educação Básica (SAEBE) e o ENEM por poder comparar diferentes grupos numa mesma escala de habilidades - processo chamado equalização e ser uma forma de avaliação mais criteriosa. Este trabalho abordou de forma básica os principais conceitos da Teoria de Resposta ao Item para ser mais uma ferramenta para tornar a TRI conhecida e assimilada na comunidade acadêmica da UFAM, tendo em vista que a TRI já está presente em nossa universidade através do processo de seleção unificada.

REFERÊNCIAS

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

ARAUJO, E. A. C; ANDRADE, D. F.; BORTOLOTTI, S. L. V. **Teoria da resposta ao item**. Ver. Esc. Enferm. USP. São Paulo, 43(Esp): 1000-8. Agosto de 2009.

Assessoria de Comunicação Social 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>. Acessado em 01 de setembro de 2010.

COSTA, M. B. F. **Técnica derivada da teoria de resposta ao item (TRI) aplicada ao setor de serviços**. Dissertação (Mestrado em Métodos Numéricos), Universidade Federal do Paraná, Paraná. 156f. 2001.

FERREIRA, F. F. G. **Escala de proficiência para o ENEM utilizando Teoria de Resposta ao Item**. Dissertação (Mestrado em Matemática e Estatística), Universidade Federal do Pará, Pará. 118f. 2009.

Ultimo Segundo. **TRI: a teoria por trás do novo Enem**. Sítio eletrônico do Portal IG. Disponível em: <<http://ultimosegundo.ig.com.br/enem/tri+a+teoria+por+tras+do+novo+enem/n1237789481686.html>>. Acessado em 20 de fevereiro de 2011.

SAMEJIMA F. A. Estimation of latent ability when responses pattern of graded scores. Psychometric Monograph. 1969

KOLEN, M. J; BRENNAN, R. L. **Test Equating – Methods and Practices**. New York: Springer 1995.

