



**UNIVERSIDADE FEDERAL DO ACRE - UFAC**  
**MESTRADO PROFISSIONAL EM MATEMÁTICA**  
**EM REDE NACIONAL - PROFMAT**



**AVALIAÇÃO EDUCACIONAL UTILIZANDO A TEORIA DE**  
**RESPOSTA AO ITEM – TRI**

**PAULO ROBERTO DE SOUZA**

**Rio Branco**  
**2014**

**PAULO ROBERTO DE SOUZA**

**AVALIAÇÃO EDUCACIONAL UTILIZANDO A TEORIA DE  
RESPOSTA AO ITEM – TRI**

Dissertação apresentada à Universidade Federal do Acre – UFAC como exigência para obtenção do título de Mestre no curso de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT.

**Orientador:** Dr. Edcarlos Miranda de Souza

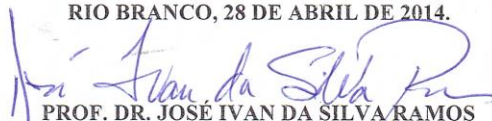
**Rio Branco  
2014**

PAULO ROBERTO DE SOUZA

AVALIAÇÃO EDUCACIONAL UTILIZANDO A TEORIA DE RESPOSTA AO ITEM – TRI

ESTE TRABALHO FOI JULGADO E APROVADO PARA A OBTENÇÃO DO TÍTULO DE MESTRE EM MATEMÁTICA NO PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL DA SOCIEDADE BRASILEIRA DE MATEMÁTICA, POLO DA UNIVERSIDADE FEDERAL DO ACRE.


RIO BRANCO, 28 DE ABRIL DE 2014.




PROF. DR. JOSÉ IVAN DA SILVA RAMOS

COORDENADOR NO POLO DA UNIVERSIDADE FEDERAL DO ACRE DO MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL – PROFMAT/UFAC


COMISSÃO EXAMINADORA



PROF. DR. EDCARLOS MIRANDA DE SOUZA (ORIENTADOR)  
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL –  
PROFMAT/UFAC



PROF. DR. MANOEL DOMINGOS FILHO (MEMBRO INTERNO)  
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL –  
PROFMAT/UFAC



PROF. DR. TOMÁS DANIEL MENÉNDEZ RODRÍGUEZ (MEMBRO EXTERNO)  
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL –  
PROFMAT/UNIR

*Dedico esse trabalho à minha família, que sempre acreditou em mim e no meu potencial.  
A meu pai José, minha mãe Olivia (IN MEMORIAN) e aos meus irmãos Sandra, Dirce e Ricardo, que sempre me apoiaram ao longo da vida em todas as minhas decisões.  
Sem eles, nada seria possível.*

## **AGRADECIMENTOS**

*Primeiramente a Deus, que dá inspiração e força para as batalhas diárias.*

*A minha família, que sempre apoiou todas as minhas decisões e acreditou nos meus sonhos.*

*Ao meu orientador, professor Edcarlos Miranda, por toda a confiança e dedicação depositada, sempre acreditando no meu potencial e me acompanhando nessa caminhada em busca do conhecimento.*

*A coordenação e professores do PROFMAT, por todo o apoio durante o curso.*

*A CAPES pelo auxílio concedido durante o período do mestrado.*

*Aos colegas de turma, sempre se ajudando ao longo dessa caminhada.*

*Ao Márcio, que foi um grande companheiro de turma, de pesquisa e de trabalho, tendo contribuído imensamente para chegar até aqui.*

*Aos meus colegas de batalha no Instituto Federal do Acre (IFAC) que sempre tinham uma palavra de apoio e ânimo nos momentos de dificuldade.*

*Aos meus amigos Maria Cristina e Fernan, mais que vizinhos, irmãos que fiz no Acre, que me ajudaram tanto em relação ao trabalho quanto nos momentos de descontração, quando precisava relaxar do trabalho.*

*Aos meus amigos de longa data Junior e Elaine, mais que irmãos, que me abriram as portas do estado do Acre e me receberam em sua casa, sempre dando apoio e sendo minha família no Acre.*

*Aos meus amigos da Universidade Federal do Acre (UFAC) e Colégio de Aplicação (CAp - UFAC) que sempre estiveram juntos torcendo e acompanhando minha batalha durante todo o curso.*

*Aos professores da banca, por todas as contribuições dadas para o enriquecimento do trabalho.*

## RESUMO

SOUZA, Paulo Roberto. **Avaliação Educacional utilizando a Teoria da Resposta ao Item – TRI**. 2014. 89 p. Dissertação (Mestrado Profissional em Matemática em Rede Nacional - PROFMAT) - Universidade Federal do Acre, Rio Branco, AC.

A Teoria de Resposta ao Item (TRI) tem sido cada vez mais utilizada para estimar de forma segura as habilidades dos indivíduos em alguma área, a partir das características dos itens que formam a avaliação. Sua aplicação mais comum são as avaliações externas em larga escala, como o Enem, o Saesp e a Prova Brasil, no intuito de classificar os respondentes ou as instituições de ensino, para as mais diversas finalidades. Este trabalho propõe o estudo desse modelo e de suas características, a partir dos referenciais teóricos, além da aplicação de um pré-teste, com a intenção de se estimar os parâmetros desses itens (característica de cada item) através de recursos computacionais, verificando quais seriam os melhores itens para a elaboração de uma avaliação escolar. Além da elaboração de um modelo de avaliação, o trabalho propõe também a discussão sobre a adoção desse modelo nas escolas, visando determinar com maior segurança, os conhecimentos adquiridos (ou desenvolvidos) pelos alunos. A partir da estimação dos parâmetros dos itens foi possível a estimação das habilidades dos respondentes e inclusive, converter esse valor para uma nota, dentro de uma escala desejada (escala do Enem) a fim de mostrar as possibilidades que se podem ter dentro da escola. Ao final da pesquisa, verificou-se que a criação de modelos de avaliação usando a TRI no ambiente escolar é viável, a partir de uma organização prévia que considere todas as etapas necessárias para o seu desenvolvimento (elaboração, calibração, aplicação e estimação de habilidades). Essa atitude pode gerar muitos benefícios ao processo ensino-aprendizagem, uma vez que as habilidades dos alunos podem ser estimadas com mais segurança, a partir de um método que leva em consideração muito mais que o total de acertos no final de uma avaliação.

**Palavras-chave:** TRI. Traço Latente. Avaliação Escolar. Função de Probabilidade. Tratamento da Informação

## ABSTRACT

SOUZA, Paulo Roberto. **Educational Evaluation using Item Response Theory -IRT**. 2014. 89 p. Dissertation (Professional Masters in Mathematics in National Network - PROFMAT) - Federal University of Acre, Rio Branco, Acre.

The Item Response Theory (IRT ) has been increasingly used to reliably estimate the skills of individuals in any area, from the characteristics of the items that form the assessment. Its most common application is external large-scale assessments such as the Enem, the Saesp and the Prova Brasil, in order to classify respondents or educational institutions for many different purposes. This work proposes the study of this model and its characteristics, based on the theoretical frameworks, in addition to applying a pre-test on "Information Processing" (on basic level), with the intention of estimating the parameters of these items (feature of each item) through computational resources, which would be checking the best items for the preparation of a school evaluation. Besides the development of an evaluation model, the paper also proposes a discussion on the adoption of this model in schools, in order to determine with greater certainty, the knowledge acquired (or developed) by students. From the estimation of item parameters was possible to estimate the abilities of the respondents and even convert this value to a grade within a desired range (scale Enem) in order to show the possibilities that it may have within the school. At the end of the survey, it was found that the creation of valuation models using the IRT in the school environment is feasible, from a previous organization that considers all the necessary steps for their development (development, calibration, application and skill estimation). This attitude can generate many benefits to the teaching-learning process, since the students' abilities can be estimated more reliably from a method that takes into account more than the total of correct answers at the end of an evaluation.

**Keywords:** IRT. Latent Trait. School Evaluation. Probability Function. Information Processing



## LISTA DE FIGURAS

<b>Figura 1</b> - Linha do tempo da TRI.....	16
<b>Figura 2</b> - CCI construída a partir do software GeoGebra .....	22
<b>Figura 3</b> - CCI com parâmetros $a$ e $c$ fixos e $b$ variando, representando itens com diferentes dificuldades, sendo $b_1 < b_2 < b_3$ .....	27
<b>Figura 4</b> - CCI com parâmetros $b$ e $c$ fixos e $a$ variando, representando itens com diferentes graus de discriminação, sendo $a_1 < a_2 < a_3$ . .....	28
<b>Figura 5</b> - CCI com parâmetros $a$ e $b$ fixos, variando o valor de $c$ , demonstrando diferentes possibilidades de acerto casual, sendo $c_1 < c_2 < c_3$ . .....	28
<b>Figura 6</b> - CCI e CII construídas juntas, fixando-se os parâmetros $a$ e $c$ . Ao parâmetro $b$ foi atribuído o valor zero ( $b = 0$ ).....	31
<b>Figura 7</b> - CCI e CII construídas juntas, fixando-se os parâmetros $a$ e $c$ . O valor do parâmetro $b$ foi aumentado, em relação a figura anterior ( $b = 3$ ).....	31
<b>Figura 8</b> - CCI e CII construídas juntas, fixando-se os parâmetros $b$ e $c$ . O parâmetro $a$ foi definido em 1,5 ( $a = 1,5$ ).....	32
<b>Figura 9</b> - CCI e CII construídas juntas, fixando-se os parâmetros $b$ e $c$ . O valor do parâmetro $a$ foi aumentado para 2,4 ( $a = 2,4$ ).....	32
<b>Figura 10</b> - CCI e CII construídas juntas, fixando-se os parâmetros $a$ e $b$ . O parâmetro $c$ foi definido em 0,2 ( $c = 0,2$ ). .....	33
<b>Figura 11</b> - CCI e CII construídas juntas, fixando-se os parâmetros $a$ e $b$ . O valor do parâmetro $c$ foi aumentado para 0,5 ( $c = 0,5$ ).....	33
<b>Figura 12</b> - Item 12 do pré-teste .....	42
<b>Figura 13</b> - Item 3 do pré-teste. ....	43
<b>Figura 14</b> - Curvas Características de todos os itens do pré-teste .....	48
<b>Figura 15</b> - Curva de Informação do Pré-teste .....	48
<b>Figura 16</b> - CCI e CII do item 1 do pré-teste.....	50
<b>Figura 17</b> - CCI e CII do item 2 do pré-teste.....	51
<b>Figura 18</b> - CCI e CII do item 3 do pré-teste.....	52
<b>Figura 19</b> - CCI e CII do item 4 do pré-teste.....	53
<b>Figura 20</b> - CCI e CII do item 6 do pré-teste.....	54
<b>Figura 21</b> - CCI e CII do item 7 do pré-teste.....	55
<b>Figura 22</b> - CCI e CII do item 8 do pré-teste.....	56
<b>Figura 23</b> - CCI e CII do item 9 do pré-teste.....	57
<b>Figura 24</b> - CCI e CII do item 10 do pré-teste.....	58
<b>Figura 25</b> - CCI e CII do item 13 do pré-teste.....	59
<b>Figura 26</b> - CCI e CII do item 16 do pré-teste.....	60
<b>Figura 27</b> - CCI e CII do item 17 do pré-teste.....	61

## LISTA DE TABELAS

<b>Tabela 1</b> - Classificação dos itens em relação à dificuldade, antes da análise com TRI. ....	43
<b>Tabela 2</b> - Escala de dificuldades utilizada como referência no trabalho.....	44
<b>Tabela 3</b> - Frequência absoluta dos acertos. ....	45
<b>Tabela 4</b> - Tabela com percentual de acertos e erros calculada pelo R. ....	46
<b>Tabela 5</b> - Parâmetros dos itens após análise feita pelo R.....	47
<b>Tabela 6</b> - Estimativa das habilidades dos respondentes na escala logística (P = participante; H = habilidade) .....	62
<b>Tabela 7</b> - Estimativa das habilidades dos respondentes na escala adotada pelo Enem (500,100).....	65

## SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 OBJETIVOS .....	12
1.2 ESTRUTURA DO TRABALHO .....	13
2 REFERENCIAL TEÓRICO.....	14
2.1 ORIGENS DA TRI.....	15
2.2 CARACTERÍSTICAS E MODELOS DA TRI .....	17
2.3 CALIBRAÇÃO.....	34
3 MATERIAIS E MÉTODOS.....	38
3.1 MATERIAIS.....	39
3.2 METODOLOGIA.....	44
4 RESULTADOS E DISCUSSÃO.....	45
4.1 ESTIMAÇÃO DOS PARÂMETROS A PARTIR DO R: CALIBRAÇÃO.....	45
4.2 ANÁLISE DOS ITENS ATRAVÉS DE SUAS CURVAS: USANDO GEOGEBRA. .	47
4.3 ESTIMAÇÃO DE HABILIDADES DOS RESPONDENTES A PARTIR DO R.....	62
4.4 A TRI DENTRO DA REALIDADE ESCOLAR: UMA POSSIBILIDADE.....	66
5 CONSIDERAÇÕES FINAIS .....	71
REFERÊNCIAS .....	73
APÊNDICE .....	75
APÊNDICE A - Pré-teste aplicado aos 216 respondentes, sobre o tema Tratamento da Informação (já excluída a questão que foi anulada). .....	76
APÊNDICE B - Tabela com os vetores de resposta do pré-teste aplicado aos 216 respondentes.....	82
APÊNDICE C - Rotina criado no Bloco de Notas para implementação da análise dos dados da Tabela pelo R. ....	88

## 1 INTRODUÇÃO

A avaliação da aprendizagem tem sido o foco de muitas discussões e pesquisas na Educação e em outras áreas relacionadas. Verificar o quanto de conhecimento foi adquirido, ou quais habilidades foram desenvolvidas pelos sujeitos tem tomado tempo e o interesse de muitos professores e pesquisadores, uma vez que, uma das principais funções da avaliação da aprendizagem é definir ações norteadoras dentro do processo educacional, indicando acertos ou possíveis mudanças na metodologia adotada.

Outra aplicação muito comum da avaliação é a classificação de indivíduos, a partir de resultados encontrados após a aplicação de algum instrumento avaliativo. Para tanto, cria-se uma escala (ou regra) que toma os resultados da avaliação (geralmente uma pontuação bruta, obtida a partir da soma dos acertos) e os converte em classificação, que em teoria, reflete os conhecimentos adquiridos (ou desenvolvidos) em determinada área. Esta aplicação é utilizada de várias maneiras, como por exemplo, o ingresso em universidades, colégios técnicos ou mesmo para o provimento de cargos no funcionalismo público, entre outras situações. Dessa forma, acredita-se que os sujeitos que conseguem êxito são aqueles com as maiores habilidades.

Entretanto, ao longo dos anos surgiram várias dúvidas em relação aos modelos de avaliação aplicados nestas situações e se eles estavam realmente refletindo o que se deseja. Seriam esses modelos capazes de avaliar e quantificar as habilidades dos alunos? Alunos que obtiveram a mesma pontuação bruta (escores brutos) tem o mesmo grau de conhecimento (habilidade)? Como medir a habilidade Matemática de um aluno, por exemplo, sendo essa uma característica que não pode ser medida diretamente, com algum instrumento ou método prático? Questionamentos como esses intrigavam os educadores e eram semelhantes a questões que já haviam surgido entre psicólogos, principalmente os psicometristas, que tinham a necessidade de medir características subjetivas de forma indireta. Essa busca gerou o que conhecemos hoje como Teoria de Resposta ao Item, numa tentativa de solucionar alguns desses problemas.

A Teoria de Resposta ao Item (TRI) é um conjunto de modelos matemáticos que tem por objetivo mensurar certas habilidades ou características de uma pessoa (que são denominadas na teoria como traços latentes), a partir das respostas dadas por ela em um conjunto de itens (teste) previamente definido. Neste modelo, analisa-se a probabilidade de ser dada uma determinada resposta, em função de certas características do item (que são

chamados de parâmetros) e de sua própria habilidade, ou seja, uma relação de dependência entre a habilidade (ou característica) e a resposta apresentada. Desse modo, quanto mais habilidade o indivíduo apresenta (ou quanto maior a característica procurada) maior será a probabilidade de ele fornecer a resposta esperada. A partir desse conjunto de respostas encontradas e de uma escala proposta, chega-se a uma medida de habilidade do indivíduo.

A utilização da TRI tem crescido bastante nos últimos tempos, principalmente na área educacional, por conta das avaliações em larga escala, que procuram verificar os níveis de aprendizagem apresentados pelos alunos ao longo da sua vida escolar ou de um determinado ano. Esses resultados têm sido usados em diversas situações, como analisar o funcionamento dos sistemas educacionais, classificar alunos para o ingresso no nível superior de ensino, classificar a qualidade de escolas a partir do desempenho de seus alunos, entre outras situações. Internacionalmente, muitas avaliações já utilizam esse modelo com resultados satisfatórios, como TOEFL (Test of English as a Foreign Language – Teste de Inglês como uma Língua Estrangeira), que é o mais conhecido exame de proficiência em língua inglesa, o SAT (Scholastic Aptitude Test ou Scholastic Assessment Test), que serve de critério para ingresso em universidades nos Estados Unidos e, as avaliações do PISA (Programa Internacional de Avaliação de Estudantes).

Assim, a TRI tem tomado um importante espaço dentro da Educação e merece uma maior atenção por parte de pesquisadores que estão ligados a essa área, ou áreas próximas, uma vez que a teoria se propõe a fazer um novo modelo de avaliação, onde o foco da análise está em cada um dos itens que formam um determinado teste, para através deles, definir as habilidades do respondente. Algumas características, como a independência entre os itens e o fato de não sofrer interferência do grupo de respondentes, tem gerado grande interesse e são responsáveis pela popularização das avaliações nesse modelo.

## **1.1 OBJETIVOS**

Assim, o objetivo deste trabalho é apresentar os conceitos da TRI e , utilizar a teoria para elaborar um modelo de avaliação, criado a partir da aplicação de um grupo de itens (teste) a uma amostra de respondentes, com o tema “Tratamento da Informação” (à nível básico), em que se pretende estimar os parâmetros dos itens e até mesmo o traço latente do grupo, utilizando o software R para efetuar as análises. Por fim, discutir sobre a aplicação deste modelo como forma de avaliação de aprendizagem dentro da escola, buscando melhorar

os métodos já existentes, e ao final, proporcionar um novo olhar sobre as metodologias de avaliação.

## **1.2 ESTRUTURA DO TRABALHO**

Com base nos objetivos, o trabalho foi dividido em 5 seções, sendo a introdução a primeira delas, trazendo um panorama geral sobre a pesquisa. A seção 2 traz o Referencial Teórico, onde será apresentado um breve levantamento histórico, com os principais pesquisadores que atuaram na criação e organização da teoria, as principais características e modelos, os métodos de calibração e sua relação com a Informática. Na seção 3, são apresentados os materiais e métodos utilizados para a realização da pesquisa. Na seção 4 estão os resultados obtidos após as análises, considerando-se as estimações de parâmetros através do R, a análise dos itens a partir das curvas no GeoGebra e a estimação de habilidades a partir do R. Além de todas essas análises, a discussão sobre a aplicação dos modelos na realidade escolar estão presentes nessa seção. A seção 5 traz as considerações finais e após isso são apresentadas as referências e o apêndice.

## 2 REFERENCIAL TEÓRICO

A TRI surge como uma alternativa diferenciada em relação aos modelos de testes tradicionais, uma vez que, em sua origem, buscava estimar as chamadas propriedades psicológicas. Araujo, Andrade e Bortolotti (2009) afirmam que estas são características individuais, não possíveis de se mensurar de forma direta, como por exemplo, o nível de qualidade de vida, potencial empreendedor, grau de satisfação, desempenho logístico, nível de estresse, entre outras situações, diferente das características físicas, como peso, altura, idade, que podem ser medidas diretamente, através de instrumentos ou mesmo de perguntas diretas.

Essas propriedades, também conhecidas como traço latente, serão estimadas a partir de itens que serão analisados através de uma escala definida previamente, que não interfere nos resultados. Conforme Andriola (2009), uma vantagem da TRI é estar focada no estudo particular de cada item que compõe o instrumento avaliativo, ou seja, “... o estudo das características métricas dos itens” (ANDRIOLA, 2009, p. 321).

Outra vantagem da TRI em relação aos métodos tradicionais é a independência de seus itens, conforme Anjos e Andrade (2012) descrevem, que “... um item mede determinado conhecimento, independente de quem o está respondendo, e a proficiência de um aluno não depende dos itens que estão sendo apresentados a ele.”(ANJOS, ANDRADE, 2012, p.1).

Essas características tem se mostrado úteis para diversas situações que vem se desenvolvendo na atualidade, como descrevem Bragion (2007), Andrade, Tavares e Valle (2000) e vários outros autores, principalmente na área das avaliações educacionais. Isso ocorre devido ao fato dos parâmetros dos itens não serem influenciados pelo grupo respondente, gerando uma maior possibilidade de estimação de traço latente. A propriedade de independência gera possibilidades como utilizar diferentes instrumentos aplicados a um mesmo grupo, que vão apresentar resultados na mesma escala de comparação ou mesmo comparar diferentes grupos que foram submetidos a instrumentos que possuem alguns itens comuns.

Para discutir sobre a utilização de uma teoria se faz necessário uma análise de sua origem, mostrando as motivações para seu desenvolvimento e os principais responsáveis por ela. Isso é importante, pois pode esclarecer as metodologias e posturas adotadas em relação ao tema na atualidade e mostrar as possíveis mudanças e adaptações que foram surgindo ao longo do tempo, na tentativa de aperfeiçoar as aplicações práticas ou teóricas que se fazem.

## 2.1 ORIGENS DA TRI

O surgimento da TRI se deu através da evolução da Psicologia, que por volta do século XIX observa um grande desenvolvimento científico. Esse movimento gera uma necessidade de se obter medidas de forma mais objetiva, para servir de embasamento para as várias pesquisas que surgiam. Essa necessidade aproxima a Psicologia das ciências exatas, como a Matemática Aplicada e a Estatística, gerando uma área conhecida como Psicometria, responsável por trabalhar com técnicas de mensuração de características ou comportamentos.

Conforme Araujo, Andrade e Bortolotti (2009), os trabalhos de Charles Spearman, no início do século XX foram responsáveis pela origem da Teoria Clássica dos Testes (TCT), importante para o desenvolvimento dos métodos operacionais utilizados pela Psicometria. Conforme Pasquali e Primi (2003), o uso dessa teoria causava certa angústia entre psicólogos, devido alguns problemas, como o fato do instrumento de medida aplicado sofrer influência do objeto a ser medido, gerando resultados dependentes dos itens utilizados e assim, causando um problema na validade e na confiança do instrumento. Outra situação era na estimação de certas características dos itens, como a dificuldade e a discriminação, que eram dependentes da amostra, e por isso, originavam resultados influenciados pelo tipo de grupo analisado. Portanto, ainda se fazia necessário à busca por uma metodologia que fosse mais independente, em que o instrumento pudesse ser aplicado a diferentes grupos sem sofrer a influência destes.

Segundo Andriola (2009) e Pasquali e Primi (2003), a busca por uma metodologia que conseguisse superar os problemas encontrados na teoria clássica seria lenta e dependeria da ação de muitos pesquisadores, em vários lugares, trazendo diferentes contribuições. A união de todos esses elementos gerou a TRI, que nunca teve a intenção de se sobrepor ou substituir a TCT, mas sim, conseguir resolver os principais problemas encontrados em relação à medida do traço latente, gerando resultados mais confiáveis.

Andriola (2009) e Pasquali e Primi (2003) trazem a cronologia desse desenvolvimento, que é apresentada resumidamente na Figura 1, com os principais nomes responsáveis e mostrando o longo processo para se chegar à metodologia atual. Além dos pesquisadores apresentados, outros também contribuíram de forma a tornar a TRI mais aceita na área da Psicologia e, posteriormente, na área da Educação, junto às avaliações. De todos os pesquisadores, dois se destacam por terem organizado as bases da teoria e tornado possível a sua popularização entre pesquisadores de várias áreas: Frederic Lord<sup>1</sup> e Allan Birnbaum<sup>2</sup>.

---

<sup>1</sup> **Frederic Lord** (1912-2000) foi psicometrista e trabalhou no ETS (Educational Testing Service), a maior



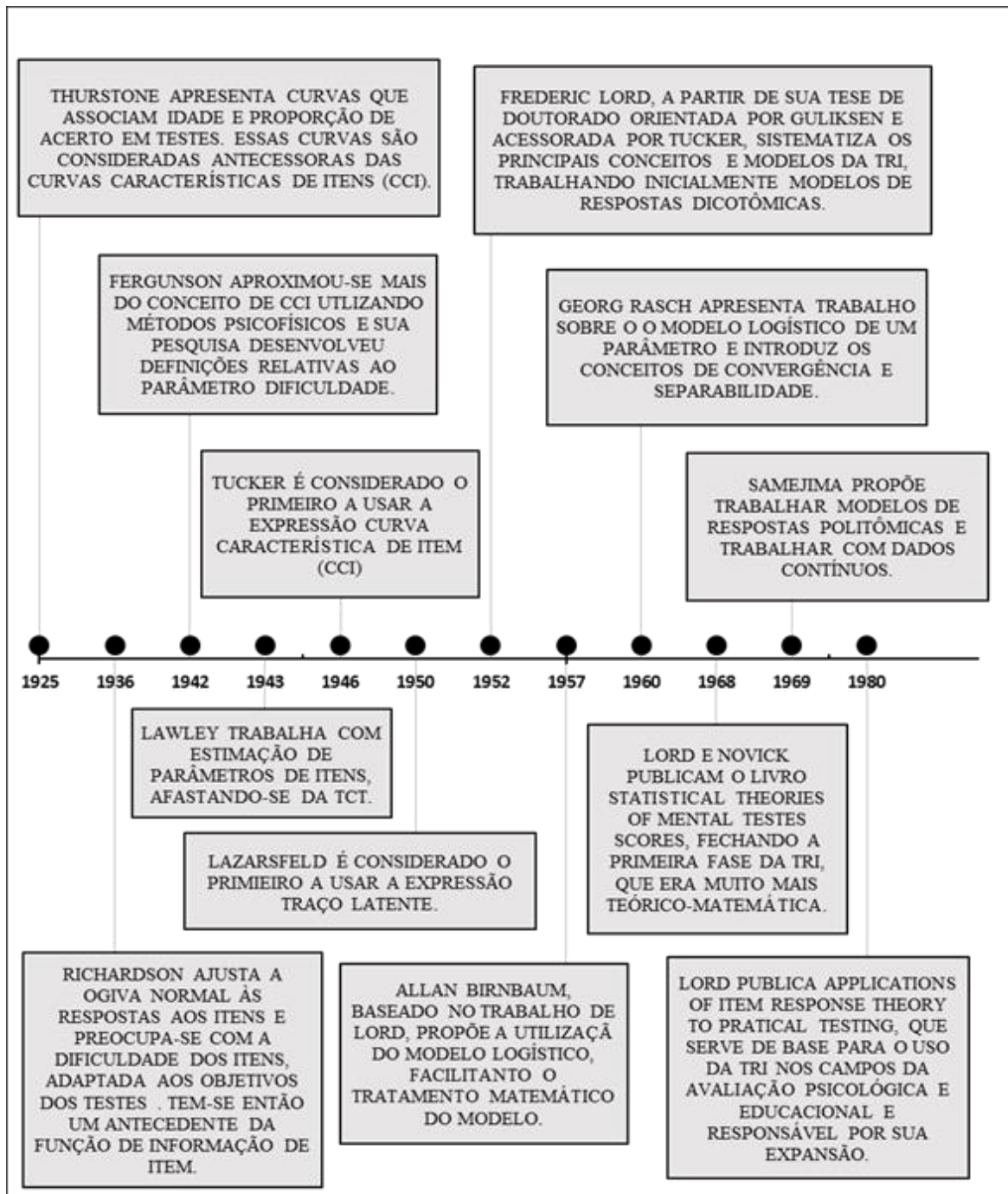


Figura 1 - Linha do tempo da TRI.

organização privada de testes e certificações educacionais do mundo (organizadora do TOEFL). A publicação de um trabalho em 1952, a partir de sua tese de doutorado é tido como o início do desenvolvimento formal da TRI e por isso é considerado por muitos como o principal responsável por sua origem.

<sup>2</sup> **Allan Birbaum** (1923-1976) foi um estatístico e professor no Courant Institute of Mathematical Sciences, na City University London e na Open University, com contribuições nas áreas de fundamentos, história, genética, inferência e psicologia, todos relacionados à estatística. Sua contribuição foi fundamental para a popularização da TRI como nova forma criar e analisar testes.

Vários autores como Pasquali e Primi (2003), Araujo, Andrade e Bortolotti (2009) e Andriola (2009) apresentam a importância do trabalho de Lord, uma vez que este propõe as bases da TRI, ao elaborar seus modelos teóricos (inicialmente o modelo unidimensional de 2 parâmetros e, posteriormente, o de 3 parâmetros) além de métodos para a estimação dos parâmetros dos itens, através da função ogiva normal. No entanto, o tratamento matemático envolvido era muito complexo e isso só seria solucionado a partir do trabalho de Birnbaum, anos mais tarde, quando foi proposta a utilização da função logística, ao invés da função ogiva normal, visto que os parâmetros aparecem de forma explícita, sem envolver processos de integração, somente trabalhando com logaritmos.

Inicialmente pode parecer estranho que, mesmo sabendo das limitações da teoria clássica, os psicometristas continuavam a usá-la. Conforme Pasquali e Primi (2003) observam, foi necessário muito estudo e desenvolvimento matemático para conseguir superar estes problemas e mesmo assim, a TRI demorou um bom tempo para ser utilizada, devido à complexidade matemática envolvida. Somente com o desenvolvimento da informática (microcomputadores e softwares), esses problemas foram superados e a metodologia passa a ser utilizada com mais frequência em diversas áreas.

## **2.2 CARACTERÍSTICAS E MODELOS DA TRI**

Ao estudar a TRI é necessário levantar as principais características e vantagens existentes ao utilizar os seus modelos para estimação de habilidades ou comportamentos. Os modelos da TRI revolucionaram a psicometria por trabalharem numa perspectiva diferente do que ocorria até então, uma vez que sua análise está focada nos itens utilizados e nas possibilidades que surgem a partir destes. Enquanto no modelo tradicional havia preocupação com os resultados totais dos testes (os chamados escores brutos ou totais), Araujo, Andrade e Bortolotti (2009) afirmam que no novo modelo a preocupação reside nos itens que compõem o teste e, através desses, geram as conclusões sobre a habilidade ou comportamento analisado.

Entre as opções de modelos de TRI destacam-se os modelos acumulativos, que serão discutidos nessa pesquisa, uma vez que despertam grande interesse na área da avaliação educacional e por isso são muito usados e citados na maioria dos trabalhos publicados, como descreve Araujo, Andrade e Bortolotti (2009). Neste modelo, a probabilidade do sujeito fornecer o resultado esperado ao item aumenta conforme o aumento de seu traço latente, gerando uma função com caráter monotônico crescente, uma vez que as duas variáveis

relacionadas apresentam o mesmo comportamento (aumentam juntas). Outro modelo de TRI existente e citado pelos autores acima é o modelo de desdobramento, que trabalha com a ideia de modelos de proximidade. Nesta situação temos que “... a probabilidade de um indivíduo dar uma resposta a um item está em função da distância entre os parâmetros do indivíduo e o de posição do item na escala...” (ARAUJO, ANDRADE E BORTOLOTTI 2009, p. 1004).

Vários autores, como Sartes e Souza-Formigoni (2013) e Andrade, Tavares e Vale (2000) têm mostrado que a opção de escolha pelo modelo de TRI a ser utilizado em determinada situação, leva em consideração alguns fatores fundamentais para o desenvolvimento da análise, sendo os principais, o número de populações envolvidas na análise, a natureza dos itens utilizados e a quantidade de habilidades ou comportamentos que se pretende estimar.

Nesta situação, será denominado população o grupo de indivíduos do qual se pretende estimar uma habilidade ou característica. Em várias situações essa população apresenta muitos sujeitos e isso inviabiliza a aplicação dos itens a todos, sendo necessária a retirada de um grupo menor (amostra) para isso. A TRI permite trabalhar com mais de um grupo dentro da mesma população ou até mesmo com populações diferentes, fazendo comparações entre resultados encontrados, que poderão estar na mesma escala, abrindo novas possibilidades de análises.

Em relação à natureza dos itens, os autores nos mostram a existência dos modelos: dicotômicos e não dicotômicos. O primeiro modelo é aquele que apresenta somente dois resultados (certo ou errado, esperado e não esperado), enquanto o segundo é aquele em que se observam mais tipos de respostas, como por exemplo, o modelo de resposta gradual, proposto por Samejima em 1969 ou o modelo de resposta nominal, proposto por Bock em 1972, em que buscavam conseguir mais informações através das respostas dos sujeitos, além do simples sim ou não. Os itens dicotômicos têm sido muito utilizados, principalmente no campo das avaliações de aprendizagem, em que o interesse é se o aluno acertou ou não a questão, desconsiderando o tipo de erro apresentado. Em algumas situações de avaliação observa-se também o processo de dicotomização de itens, onde um item não dicotômico passará a ser tratado como dicotômico, procurando dessa forma facilitar o trabalho com esse item.

Sobre as hipóteses básicas citadas pelos autores temos a quantidade de traços latentes analisados pelos modelos, havendo duas situações possíveis: se o modelo analisa um único traço latente (conhecido como modelo unidimensional) ou se o modelo analisa mais de um traço (modelo multidimensional). O modelo unidimensional tem sido o mais explorado e será o foco dessa pesquisa, uma vez que é o mais utilizado no campo das avaliações de

aprendizagem. Além disso, como nos traz Pasquali e Primi (2003), os pesquisadores na área de Psicologia concordam que os resultados apresentados por um sujeito podem ser influenciados por várias habilidades, mas consideram que existe uma que seja dominante e principal responsável de desempenho do sujeito no teste (chamado traço dominante). Assim, acredita-se que esse traço dominante seja o medido no teste, garantindo a unidimensionalidade do modelo.

A garantia da unidimensionalidade do modelo implica na outra hipótese, conhecida como independência local ou invariância. A invariância garante que, dada certa habilidade, as respostas fornecidas aos diferentes itens do instrumento são independentes entre si. Isso é de grande importância no momento de se estimar os parâmetros do modelo.

Os parâmetros dos itens são outra característica importante da TRI e merecem a nossa atenção. Eles são características particulares de cada item, que através de funções matemáticas, relacionam o traço latente do indivíduo (uma habilidade ou comportamento) com a probabilidade dele fornecer a resposta esperada. A estimação desses parâmetros é parte fundamental das análises da TRI pois a partir dos valores encontrados será possível a estimação do traço latente. Os modelos de TRI mais comuns apresentam um, dois ou três parâmetros, cuja estimação pode ser feita através de diferentes modelos estatísticos, utilizando softwares que auxiliam nos cálculos. Os modelos costumam apresentar os seguintes parâmetros:

- **Parâmetro  $b$** , definido como grau de dificuldade, representa a habilidade necessária ao indivíduo para responder o item. Desse modo, se um indivíduo necessita de muita habilidade para responder um item é provável este tenha alto grau de dificuldade. O inverso também é válido, ou seja, itens com pouca dificuldade requerem baixa habilidade do respondente. Por sua natureza, este parâmetro deve ser medido na mesma escala que a habilidade do sujeito, que será definida conforme a situação que está sendo avaliada, podendo assumir qualquer valor entre os números reais. Este parâmetro está presente nos principais modelos de TRI;
- **Parâmetro  $a$** , definido como grau de discriminação, representa a capacidade do item de diferenciar o grau de habilidade procurada entre os respondentes. Nesse caso, um item com baixo grau de discriminação costuma fazer com que pessoas de diferentes graus de habilidade tenham probabilidades de acertar próximas, atrapalhando a estimação do traço latente. Quando este parâmetro é alto, o item conseguirá diferenciar pessoas que tem habilidade imediatamente superior ao nível de dificuldade do mesmo, daquelas que tem habilidade imediatamente inferior à dificuldade do item. É desejável

que um bom item tenha um alto grau de discriminação. Bragion (2007), afirma que quanto maior o valor desse parâmetro, maior será a sensibilidade do modelo em relação às variações de habilidade, na região próxima ao ponto de dificuldade. Este parâmetro aparece nos modelos de dois e três parâmetros sendo representado por números reais positivos;

- **Parâmetro  $c$** , definido como probabilidade de acerto casual, aparece nos modelos de três parâmetros com itens de respostas dicotômicas, representando a probabilidade de um indivíduo fornecer o resultado esperado, mesmo com pouca ou sem ter a habilidade necessária para isso. Quando um sujeito possui certa habilidade, ele responde os itens com dificuldade menor que ela, e continua assim até que a dificuldade dos itens supere sua habilidade, formando um certo padrão contínuo de resultados. No entanto, se o sujeito apresenta acertos muito variados (erra questões muito fáceis, acertas as difíceis) formando um padrão aleatório de acertos, supõe-se que exista o acerto casual. Quanto mais isso ocorre, maior será o valor desse parâmetro, que sendo uma probabilidade, será um valor entre 0 e 1.

Um bom teste é aquele que consegue ter itens em vários graus da escala de dificuldade, ou seja, com diversos itens e estes percorrendo toda a escala. Além disso, é desejado que todos estes itens tenham alto poder de discriminação. Por exemplo, um item pode ser fácil com alto poder de discriminação entre indivíduos com baixa habilidade, ou um item médio com alto poder de discriminação entre indivíduos com média habilidade e assim por diante.

Bragion (2007) salienta uma importante propriedade da TRI, referente a invariância de seus parâmetros em relação às habilidades dos sujeitos analisados. Os parâmetros interferem na probabilidade do sujeito fornecer a resposta esperada, mas não dependem do grau do traço latente e, nem este é afetado pelos valores utilizados para determiná-lo, mostrando grande vantagem na utilização desses modelos.

Os modelos de TRI vão relacionar esses parâmetros ao traço latente, através da probabilidade dos sujeitos fornecerem as respostas esperadas (respostas corretas no caso de avaliações de aprendizagem). O modelo unidimensional de três parâmetros é o mais completo dessa categoria, inclusive podendo ser aplicado às situações em que estão envolvidos menos parâmetros. Nesta situação, o parâmetro  $c$  assume o valor 0 (quando não se considera a possibilidade de acerto casual) e o parâmetro  $a$  assume o valor 1 (quando não há discriminação no item), podendo dessa forma chegar ao modelo de um ou dois parâmetros. O modelo matemático que representa essa relação entre parâmetros e traço latente é dado pela

função logística, adotado por Birnbau e Rasch ainda na década de 1960, de manipulação mais prática em relação à função ogiva normal, proposta por Lord. O Modelo Logístico de 3 parâmetros (ML3), conhecido como Função de Resposta do Item (FRI), é dado por:

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \cdot \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}},$$

com  $i=1, 2, 3, \dots, I$ , representando os itens propostos para avaliar o traço latente;

$j = 1, 2, 3, \dots, n$ , representando os indivíduos da amostra que serão analisados;

$U_{ij}$  é uma variável dicotômica que representa a resposta do indivíduo  $j$  ao item  $i$ . O valor de  $U_{ij}$  será 1 se o indivíduo der a resposta esperada ao item (ou satisfizer suas condições) e 0 caso isso não ocorra;

$\theta_j$  representa o traço latente (habilidade) do respondente  $j$ ;

$P(U_{ij} = 1|\theta_j)$  é a probabilidade do respondente  $j$ , a partir de seu traço latente, fornecer a resposta esperada ao item  $i$  (ou satisfizer suas condições);

$b_i$  é o parâmetro dificuldade, medido na mesma escala da habilidade  $\theta_j$ ;

$a_i$  é o parâmetro discriminação;

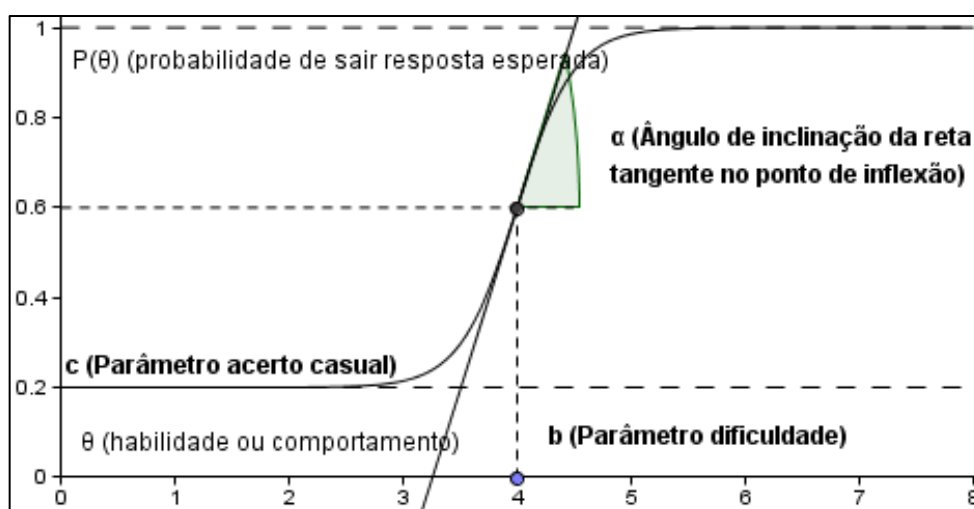
$c_i$  é o parâmetro do acerto casual;

$D$  é uma constante que vale 1, mas pode assumir o valor 1,7 quando se deseja que os valores do modelo se aproximem da função ogiva normal.

Através desse modelo, conforme Sartes e Souza-Formigoni (2013), a probabilidade de um indivíduo responder um item da forma esperada (ou satisfizer as características deste) será analisada em função de seu traço latente, sendo seu resultado afetado pelos parâmetros dificuldade ( $b$ ), discriminação ( $a$ ) e acerto casual ( $c$ ).

Com base nesse modelo será construída a Curva Característica do Item (CCI), em que estarão representadas graficamente as suas informações, através de uma função de probabilidade particular de cada item, com um formato de sigmoide. Essa curva será construída num plano cartesiano em que o eixo das abscissas representa o grau de habilidade ou comportamento do sujeito ( $\theta$ ), variando conforme escala determinada pela situação que está sendo avaliada e onde serão mais importantes as relações de ordem entre os elementos do eixo que propriamente os valores destes. O eixo das ordenadas representa a probabilidade do indivíduo apresentar a resposta esperada pelo item, variando então de 0 a 1. O comportamento

da curva é monotônico crescente quando representa modelos acumulativos e, um formato de sino com um pico, quando representa os modelos de desdobramento. Na situação envolvendo modelos acumulativos temos que, quanto maior o traço latente do sujeito analisado, maior será a probabilidade de ele fornecer resposta esperada, ficando clara a relação estabelecida entre a probabilidade de resposta esperada e a intensidade do traço latente procurado. É claro que por maior que seja a habilidade do indivíduo, existe uma assíntota para esta probabilidade no valor igual a 1 e a partir de uma certa habilidade as probabilidades dos indivíduos acertarem o item estará convergindo para 1. É razoável que nenhum indivíduo alcance os valores da assíntota, pois por maior que seja a habilidade do mesmo, sempre existirá uma pequena probabilidade do mesmo errar o item, da mesma forma, por menor que seja a habilidade de um indivíduo, a probabilidade do mesmo acertar o item jamais alcançará o valor zero, pois sempre haverá a possibilidade do mesmo acertar por acaso. Na Figura 2 a seguir observa-se um exemplo de CCI construída com o auxílio do software GeoGebra<sup>3</sup>, a partir do modelo de 3 parâmetros, em que foram atribuídos valores para os parâmetros  $a$ ,  $b$  e  $c$  ( $a = 4$ ;  $b = 4$ ;  $c = 0,2$ ).



**Figura 2** - CCI construída a partir do software GeoGebra.

Sendo a CCI uma função que relaciona a probabilidade do sujeito dar a resposta esperada (variável dependente da função) com seu traço latente (variável independente da função), pode-se concluir que o conjunto domínio dessa função será o conjunto dos números

<sup>3</sup> No site [www.geogebra.org](http://www.geogebra.org) encontra-se a descrição sobre o software: “O GeoGebra é um software de matemática dinâmica gratuito e multi-plataforma para todos os níveis de ensino, que combina geometria, álgebra, tabelas, gráficos, estatística e cálculo em um único sistema.”

reais ( $\text{Domínio} (P(U_{ij} = 1|\theta_j)) = \mathbb{R}$ ) uma vez que os valores da habilidade serão uma escala definida conforme a situação que esteja sendo avaliada, podendo assumir valores positivos ou negativos. Já a imagem dessa função será um intervalo aberto, iniciando no valor do parâmetro  $c$  (probabilidade de acerto casual), que ocorre no caso de um indivíduo não ter o traço latente e ainda assim fornecer o resultado esperado, variando até  $1$ , ( $\text{Imagem} (P(U_{ij} = 1|\theta_j)) = (c, 1)$ ). Se o modelo adotado não utiliza o parâmetro de acerto casual ( $c = 0$ ), a imagem dessa função será o intervalo aberto de extremos  $0$  e  $1$ .

É possível notar que os parâmetros do item estarão presentes de forma marcante na CCI, influenciando seu formato e comportamento, independente se trabalhamos com modelos de 1, 2 ou 3 parâmetros. Além de ser um dos extremos do conjunto imagem da função, o parâmetro  $c_i$  também será a assíntota horizontal inferior desta curva ( $y = c_i$ ), assim como o ponto 1 será sua assíntota horizontal superior ( $y = 1$ ). Isso é possível de se verificar quando analisamos o comportamento dessa função para os valores de  $\theta$  variando para  $-\infty$  e para  $+\infty$ , ou seja, o limite da função no infinito. Observe o comportamento dessa função quando se faz  $\theta$  tender a infinito pela esquerda.

$$\begin{aligned} \lim_{\theta \rightarrow -\infty} (P(U_{ij} = 1|\theta_j)) &= \lim_{\theta \rightarrow -\infty} \left( c_i + (1 - c_i) \cdot \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) = \\ &= c_i + (1 - c_i) \cdot \lim_{\theta \rightarrow -\infty} \left( \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \end{aligned}$$

Através das propriedades dos limites sabe-se que o limite da soma é igual à soma dos limites e, o limite de um produto é o produto dos limites. Na situação acima, como  $c_i$  é constante, o limite dele em relação à  $\theta$  será o próprio, reduzindo o problema a calcular o limite da fração na última parte da expressão, onde aparece a variável. Com o valor de  $\theta$  indo para infinito à esquerda (negativo) o conteúdo dos parênteses na exponencial será negativo, mas ao ser multiplicado por  $-Da_i$  ficará positivo e a exponencial tenderá ao infinito com valores positivos (o valor  $1$  somado a exponencial não afetará o resultado, que continuará tendendo a infinito). Como essa exponencial é denominador de uma fração e sabe-se que, quanto maior o denominador menor será o seu valor, no infinito essa fração tenderá a zero. Logo se tem:

$$\lim_{\theta \rightarrow -\infty} (P(U_{ij} = 1|\theta_j)) = c_i + (1 - c_i) \cdot 0 = c_i$$



Portanto, a reta  $y = c_i$  é assíntota horizontal inferior da CCI, como podemos observar na Figura 2, onde o parâmetro  $c_i$  assumiu o valor 0,2 (a reta  $y = 0,2$  é a assíntota horizontal inferior da curva). Caso o modelo utilizado seja de 1 ou 2 parâmetros ( $c_i = 0$ ), a assíntota inferior será o próprio eixo das abscissas (eixo das habilidades).

Para determinar a assíntota vertical superior utiliza-se o mesmo processo acima, valendo-se das propriedades e manipulações algébricas já utilizadas. A diferença está em se fazer a variável  $\theta$  tender a infinito pela direita ( $+\infty$ ). Dessa forma tem-se:

$$\begin{aligned} \lim_{\theta \rightarrow +\infty} \left( P(U_{ij} = 1 | \theta_j) \right) &= \lim_{\theta \rightarrow +\infty} \left( c_i + (1 - c_i) \cdot \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) = \\ &= c_i + (1 - c_i) \cdot \lim_{\theta \rightarrow +\infty} \left( \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \end{aligned}$$

Ao calcular o limite do denominador da fração, como  $\theta$  tende a  $+\infty$ , a expressão nos parênteses da exponencial ficará positiva (o valor de  $b_i$  não afetará esse comportamento). Ao multiplicar essa expressão por  $-Da_i$  ela se tornará negativa e a exponencial tenderá ao infinito pela esquerda, quando os valores dela vão ficando cada vez menores, aproximando-se de 0 no limite. Tem-se:

$$\lim_{\theta \rightarrow +\infty} \left( P(U_{ij} = 1 | \theta_j) \right) = c_i + (1 - c_i) \cdot \frac{1}{1 + 0} = c_i + 1 - c_i = 1$$

Assim observa-se que a reta  $y = 1$  é assíntota horizontal superior dessa curva e isso independe da quantidade de parâmetros utilizados ou modelo adotado, uma vez que essa função é uma probabilidade e naturalmente seus resultados não iriam ultrapassar 1.

Ao trabalhar com modelos dicotômicos de TRI, busca-se analisar qual a probabilidade de um sujeito dar a resposta esperada em determinado item. É importante notar que nesse tipo de item só existe uma resposta esperada (“resposta correta”). No entanto, essa probabilidade será afetada pela habilidade ou comportamento particular do respondente (o chamado traço latente), uma vez que, se ele possui muita, ou pouca habilidade, esta afetaria diretamente sua probabilidade de resposta (sujeitos com pouca habilidade teriam menor probabilidade de dar a resposta esperada, enquanto que os que possuem alta habilidade teriam probabilidade maior).

No modelo de três parâmetros as características são similares às aquelas apresentadas para os modelos de 2 parâmetros, mas deve-se levar em consideração o acerto casual, que também irá influenciar a probabilidade. Nesse caso, quando o sujeito apresenta grau de traço

latente igual à dificuldade da questão, terá a probabilidade da resposta esperada igual à média entre os extremos de seu conjunto imagem (ponto médio do conjunto imagem), ou seja,  $\frac{1+c}{2}$ . Este resultado pode ser aplicado também ao modelo de dois parâmetros, quando assumimos  $c = 0$ , em que se chegará à probabilidade 0,5. Esses resultados podem ser observados nas CCI e encontrados através de manipulações no modelo matemático, fazendo  $\theta_j = b_i$ . Observe:

$$\begin{aligned} P(U_{ij} = 1 | \theta_j = b_i) &= c_i + (1 - c_i) \cdot \frac{1}{1 + e^{-Da_i(b_i - b_i)}} \\ &= c_i + (1 - c_i) \cdot \frac{1}{1 + e^0} = c_i + \frac{(1 - c_i)}{2} = \frac{1 + c_i}{2} \end{aligned}$$

Tomando o exemplo da Figura 2, o ponto médio do conjunto imagem se encontra no valor 0,6, pois seu parâmetro de acerto casual é 0,2. Traçando um segmento de reta paralelo ao eixo das habilidades, passando por 0,6, ele intercepta a CCI no ponto cuja abscissa tem valor 4, que é o valor do parâmetro dificuldade. Se o acerto casual desse exemplo fosse descartado, esse gráfico seria transladado verticalmente para baixo e a ordenada desse ponto mudaria de 0,6 para 0,5, que seria o ponto médio do novo conjunto imagem da função.

Além das situações apresentadas anteriormente, o parâmetro  $b$  ainda apresenta outra importante participação na CCI, pois ele será a abscissa do ponto de inflexão dessa curva (ponto em que a curva muda sua concavidade). Isso pode ser demonstrado através do processo de derivação da função  $P(U_{ij} = 1 | \theta_j)$ , uma vez que a raiz da segunda derivada da função é a abscissa do ponto de inflexão. Segue abaixo a demonstração dessa afirmação, lembrando que os parâmetros  $a$ ,  $b$  e  $c$  são diferentes de zero e  $D$  é uma constante igual a 1 (ou 1,7).

$$\begin{aligned} \frac{dP}{d\theta} &= (1 - c_i) \cdot \frac{Da_i e^{-Da_i(\theta_j - b_i)}}{(1 + e^{-Da_i(\theta_j - b_i)})^2} \\ \frac{d^2P}{d\theta^2} &= D^2 a_i^2 \cdot e^{-Da_i(\theta_j - b_i)} \cdot (1 - c_i) \cdot \frac{e^{-Da_i(\theta_j - b_i)} - 1}{(1 + e^{-Da_i(\theta_j - b_i)})^3} \\ \frac{d^2P}{d\theta^2} = 0 &\Rightarrow e^{-Da_i(\theta_j - b_i)} - 1 = 0 \Rightarrow e^{-Da_i(\theta_j - b_i)} = 1 \Rightarrow -Da_i(\theta_j - b_i) = 0 \Rightarrow \\ &\theta_j - b_i = 0 \Rightarrow \theta_j = b_i \rightarrow \text{ponto de inflexão} \end{aligned}$$

O parâmetro  $a$  (grau de discriminação) também estará presente na CCI, através da

inclinação da curva em seu ponto de inflexão. Esta inclinação é encontrada através da reta tangente à curva no ponto de inflexão e será proporcional ao parâmetro  $a$  (lembrando que  $D$  é constante e vale 1), podendo ser calculada a partir da derivada de  $P(U_{ij} = 1|\theta_j)$  no ponto  $\theta_j = b_i$ . Assim:

$$\frac{dP}{d\theta} = (1 - c_i) \cdot \frac{Da_i e^{-Da_i(\theta_j - b_i)}}{(1 + e^{-Da_i(\theta_j - b_i)})^2}$$

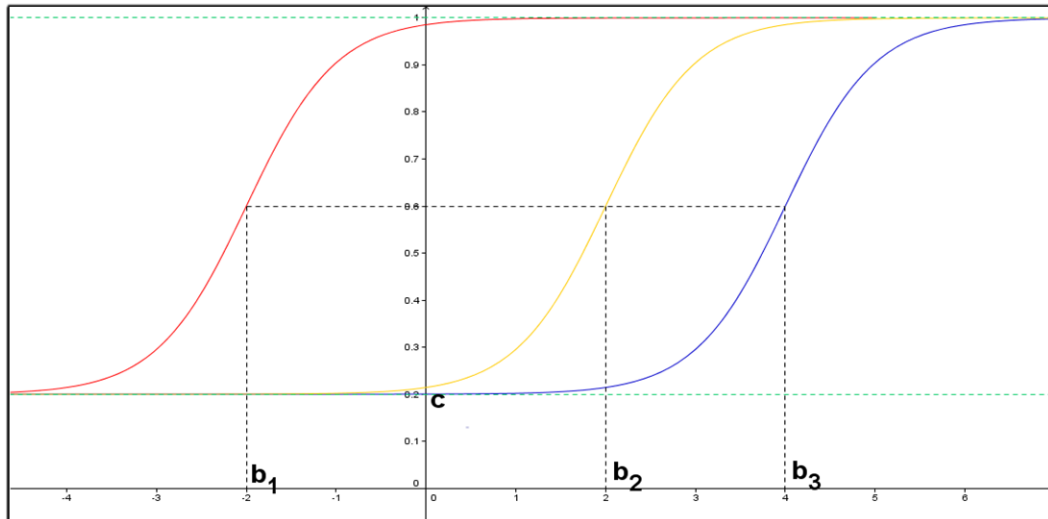
$$\frac{dP(b)}{d\theta} = (1 - c_i) \cdot \frac{Da_i e^{-Da_i(b_i - b_i)}}{(1 + e^{-Da_i(b_i - b_i)})^2} = (1 - c_i) \cdot \frac{Da_i e^0}{(1 + e^0)^2} = (1 - c_i) \cdot \frac{Da_i}{2^2} = \frac{(1 - c_i)}{4} \cdot a_i$$

Dessa forma observa-se que o grau de discriminação do item está diretamente relacionado com a inclinação da CCI (a inclinação da reta tangente no ponto de inflexão é diretamente proporcional ao parâmetro  $a$ ), de modo que, quanto maior a discriminação do item, mais inclinada será a curva. A partir disso, pode-se afirmar que o parâmetro  $a$  não poderá apresentar valores negativos, pois uma vez que isso ocorresse, significaria que a reta tangente à curva na inflexão seria decrescente e, conseqüentemente, à medida que aumentasse a habilidade, a probabilidade de se encontrar a resposta esperada diminuiria.

Para ilustrar melhor toda a discussão apresentada anteriormente, apresenta-se a seguir várias representações de CCI, onde se fixou dois parâmetros e variou-se o terceiro, ficando mais claro como isso interfere no comportamento do modelo.

- **VARIANDO O PARÂMETRO  $b$  ( $a$  e  $c$  fixos)**

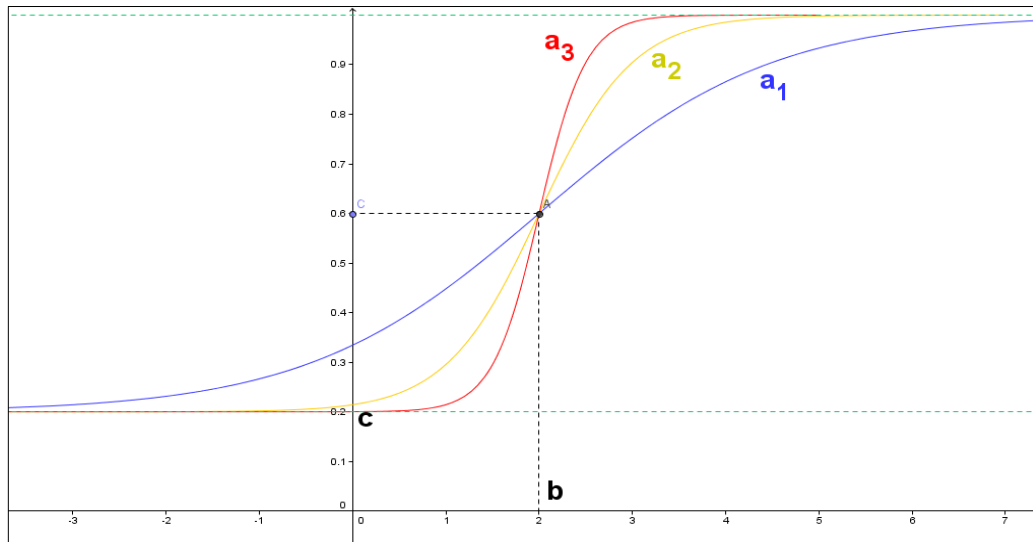
Na Figura 3 observamos modelos de CCI em que foram fixados os parâmetros  $a$  e  $c$ , variando-se o parâmetro  $b$  (parâmetro de dificuldade). Quanto maior a dificuldade do item mais deslocado a direita ficará a curva e maior deverá ser a habilidade (ou comportamento) do respondente para que consiga apresentar a resposta esperada. A escala de habilidades será um intervalo numérico real, definido conforme a situação que esteja sendo avaliada pelo modelo.



**Figura 3** - CCI com parâmetros  $a$  e  $c$  fixos e  $b$  variando, representando itens com diferentes dificuldades, sendo  $b_1 < b_2 < b_3$ .

- **VARIANDO O PARÂMETRO  $a$  ( $b$  e  $c$  fixos)**

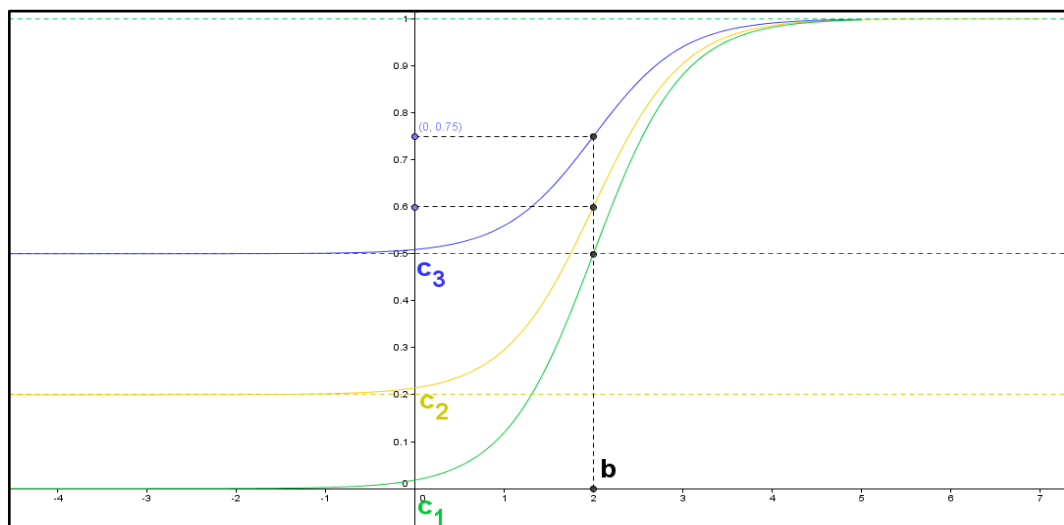
Na Figura 4 observamos modelos de CCI em que foram fixados os parâmetros  $b$  e  $c$ , variando-se o parâmetro  $a$ . Quanto maior o grau de discriminação do item mais inclinada ficará a curva (aproximando-se de uma inclinação de  $90^\circ$  em relação ao eixo das habilidades) dividindo os respondentes entre os que possuem habilidade menor que a dificuldade do item e os que possuem habilidade maior que a dificuldade. O item com parâmetro  $a_1$  possui baixa discriminação em relação aos outros e observa-se na curva que, na região próxima ao valor  $b$ , não existe grande variação na probabilidade do respondente fornecer a resposta esperada, ou seja, pessoas com diferentes habilidades tem a probabilidade muito próxima de fornecer a resposta correta. Já o item com parâmetro  $a_3$ , o maior entre eles, possui alta discriminação e vemos uma variação maior da probabilidade na região próxima ao valor de  $b$ , discriminando melhor a habilidade do respondente.



**Figura 4** - CCI com parâmetros  $b$  e  $c$  fixos e  $a$  variando, representando itens com diferentes graus de discriminação, sendo  $a_1 < a_2 < a_3$ .

- **VARIANDO O PARÂMETRO  $c$  ( $a$  e  $b$  fixos)**

Na Figura 5 observamos os modelos de CCI em que foram fixados os parâmetros  $a$  e  $b$  e variou-se o parâmetro  $c$ . Como esse parâmetro representa a probabilidade do acerto casual, ele terá um comportamento assintótico na curva, sendo a menor probabilidade do respondente fornecer a resposta esperada. O valor de  $c$ , interfere na probabilidade de um indivíduo que possui habilidade igual à dificuldade do item fornecer a resposta esperada. À medida que o valor de  $c$  aumenta, a curva é deslocada para cima, ficando mais achatada entre suas assíntotas. A curva de parâmetro  $c_1$  representa uma CCI no modelo de 2 parâmetros, pois  $c_1 = 0$ .



**Figura 5** - CCI com parâmetros  $a$  e  $b$  fixos, variando o valor de  $c$ , demonstrando diferentes possibilidades de acerto casual, sendo  $c_1 < c_2 < c_3$ .

Além da CCI, outra importante medida relacionada ao item é a chamada Função de Informação do item (FII), que segundo Andrade, Tavares e Vale (2000), analisa a quantidade de informação contida em um item dentro da escala de habilidade do respondente. Esta função é dada por:

$$I_i(\theta) = \frac{\left[\frac{d}{d\theta}P_i(\theta)\right]^2}{P_i(\theta) \cdot Q_i(\theta)}, \text{ em que:}$$

$I_i(\theta)$  é a informação fornecida pelo item  $i$  no nível de habilidade  $\theta$ ;

$P_i(\theta) = P(U_{ij} = 1|\theta_j)$ , ou seja, a probabilidade do item ter a resposta esperada em função da habilidade;

$Q_i(\theta) = 1 - P_i(\theta)$ , ou seja, o complementar da probabilidade  $P_i(\theta)$ .

Sabendo que a derivada de  $P_i(\theta)$  é  $\frac{d}{d\theta}P_i(\theta) = (1 - c_i) \cdot \frac{Da_i e^{-Da_i(\theta_j - b_i)}}{(1 + e^{-Da_i(\theta_j - b_i)})^2}$  e a usando

na função de informação do item  $I_i(\theta)$ , temos que:

$$\begin{aligned} I_i(\theta) &= \frac{\left[\frac{d}{d\theta}P_i(\theta)\right]^2}{P_i(\theta) \cdot Q_i(\theta)} = \frac{\frac{D^2 a_i^2 \cdot (1 - c_i)^2 \cdot (e^{-Da_i(\theta_j - b_i)})^2}{(1 + e^{-Da_i(\theta_j - b_i)})^4}}{\left[c_i + \frac{(1 - c_i)}{1 + e^{-Da_i(\theta_j - b_i)}}\right] \cdot \left[1 - c_i + \frac{(1 - c_i)}{1 + e^{-Da_i(\theta_j - b_i)}}\right]} = \\ &= \frac{\frac{D^2 a_i^2 \cdot (1 - c_i)^2 \cdot (e^{-Da_i(\theta_j - b_i)})^2}{(1 + e^{-Da_i(\theta_j - b_i)})^4}}{\left[\frac{c_i(1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)}{1 + e^{-Da_i(\theta_j - b_i)}}\right] \cdot \left[\frac{(1 - c_i) \cdot (1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)}{1 + e^{-Da_i(\theta_j - b_i)}}\right]} = \\ &= \frac{\frac{D^2 a_i^2 \cdot (1 - c_i)^2 \cdot (e^{-Da_i(\theta_j - b_i)})^2}{(1 + e^{-Da_i(\theta_j - b_i)})^4}}{\left\{\frac{\left[c_i(1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)\right] \cdot (1 - c_i) \cdot e^{-Da_i(\theta_j - b_i)}}{\left[1 + e^{-Da_i(\theta_j - b_i)}\right]^2}\right\}} \end{aligned}$$

$$\begin{aligned}
&= \frac{D^2 a_i^2 \cdot (1 - c_i)^2 \cdot (e^{-Da_i(\theta_j - b_i)})^2}{(1 + e^{-Da_i(\theta_j - b_i)})^4} \cdot \frac{[1 + e^{-Da_i(\theta_j - b_i)}]^2}{[c_i(1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)] \cdot (1 - c_i) \cdot e^{-Da_i(\theta_j - b_i)}} = \\
&= D^2 a_i^2 \cdot \frac{(1 - c_i) \cdot e^{-Da_i(\theta_j - b_i)}}{1 + e^{-Da_i(\theta_j - b_i)}} \cdot \frac{1}{[c_i(1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)] \cdot (1 + e^{-Da_i(\theta_j - b_i)})} = \\
&= D^2 a_i^2 \cdot Q(\theta) \cdot \frac{1}{\frac{[c_i(1 + e^{-Da_i(\theta_j - b_i)}) + (1 - c_i)] \cdot (1 + e^{-Da_i(\theta_j - b_i)})^2}{1 + e^{-Da_i(\theta_j - b_i)}}} = \\
&= D^2 a_i^2 \cdot Q(\theta) \cdot \frac{1}{P(\theta)} \cdot \frac{1}{(1 + e^{-Da_i(\theta_j - b_i)})^2} = D^2 a_i^2 \cdot \frac{Q_i(\theta)}{P_i(\theta)} \cdot \left[ \frac{P_i(\theta) - c_i}{1 - c_i} \right]^2
\end{aligned}$$

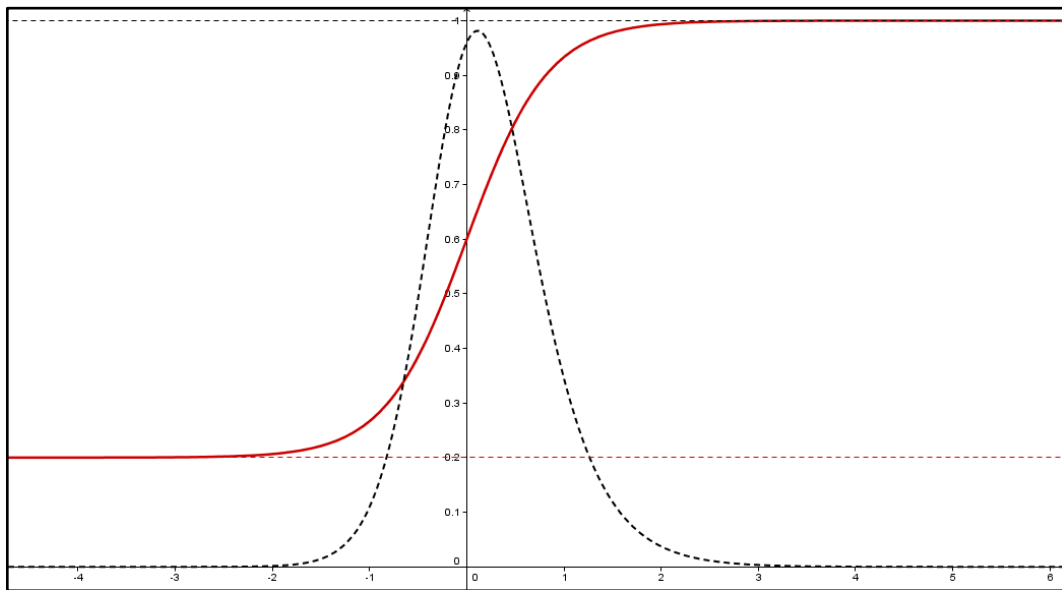
Conforme Andrade, Tavares e Vale (2000) descrevem, através dessa função pode-se verificar como os parâmetros são importantes em relação à informação trazida pelo item. Um bom item, que contribui com muita informação para o valor de  $\theta$ , é aquele em que há um baixo valor de acerto casual (parâmetro  $c$  aproximando-se de zero), alta discriminação (parâmetro  $a$  grande) e dificuldade igual à habilidade (parâmetro  $b$  aproximando-se do valor de  $\theta$ ).

A função de informação do item depende da probabilidade de ser dada a resposta esperada em determinado item e esta probabilidade depende da habilidade do respondente. Logo, essa função depende do valor de  $\theta$  e pode ser representada graficamente, da mesma forma como foi feita com a CCI, utilizando eixos que relacionem habilidade do respondente e probabilidade de fornecer a resposta esperada. Esta curva é conhecida como Curva de Informação do Item (CII) e em conjunto com a CCI pode demonstrar importantes características dos itens, através da variação dos parâmetros das funções. Nas Figuras 6 a 11, serão apresentadas CII acompanhadas das respectivas CCI, para que se possam verificar seus comportamentos.

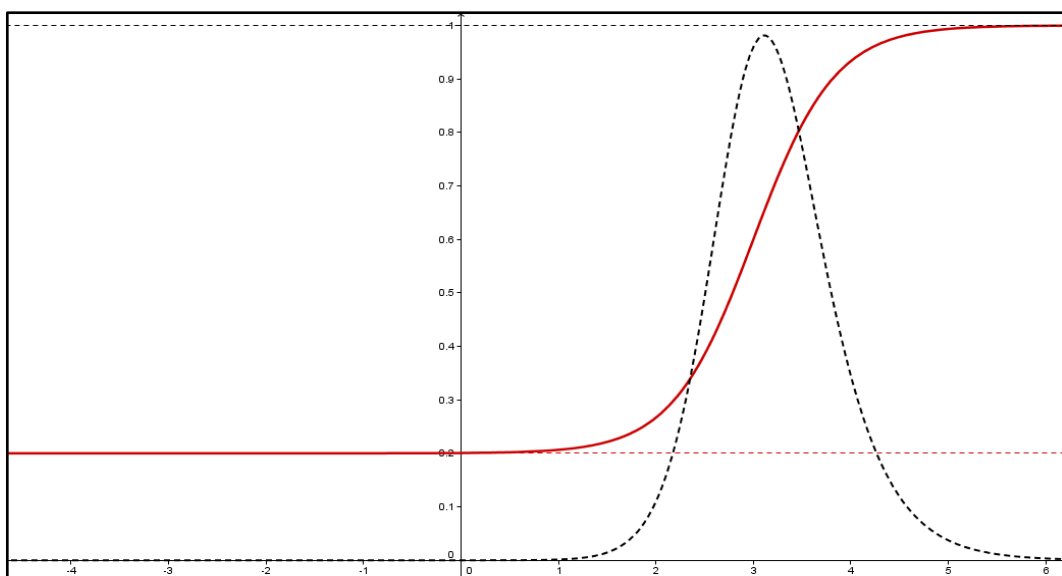
- **VARIANDO O PARÂMETRO  $b$  ( $a$  e  $c$  fixos)**

A variação do parâmetro  $b$  dentro dos modelos matemáticos da CCI causa mudança de posição da curva em relação ao eixo das habilidades, ou seja, questões com maior dificuldade

necessitam de mais habilidade para serem respondidas e a curva vai sendo deslocada para direita. Nas Figuras 6 e 7, os parâmetros  $a$  e  $c$  estão fixos ( $a = 2,4$  e  $c = 0,2$ ) e  $b$  variou (Figura 6,  $b = 0$ ; Figura 7,  $b = 3$ ), sendo que a CII atinge seu máximo próximo ao parâmetro  $b$  (curva tracejada). No momento em que o grau de dificuldade aumenta tanto a CCI quanto a CII são deslocadas para direita no eixo das habilidades (Figura 7), indicando a necessidade de maior habilidade para responder ao item proposto.



**Figura 6** - CCI e CII construídas juntas, fixando-se os parâmetros  $a$  e  $c$ . Ao parâmetro  $b$  foi atribuído o valor zero ( $b = 0$ ).

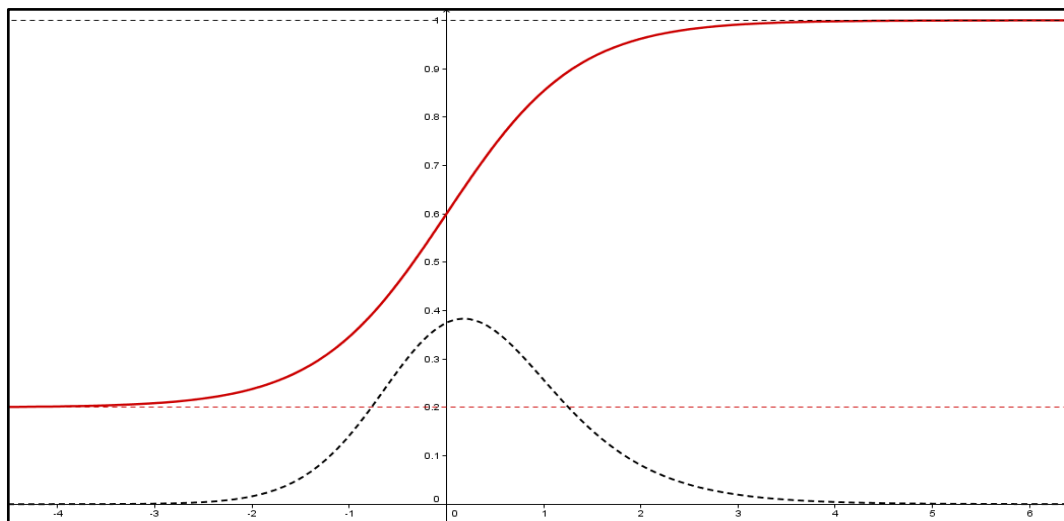


**Figura 7** - CCI e CII construídas juntas, fixando-se os parâmetros  $a$  e  $c$ . O valor do parâmetro  $b$  foi aumentado, em relação a figura anterior ( $b = 3$ ).

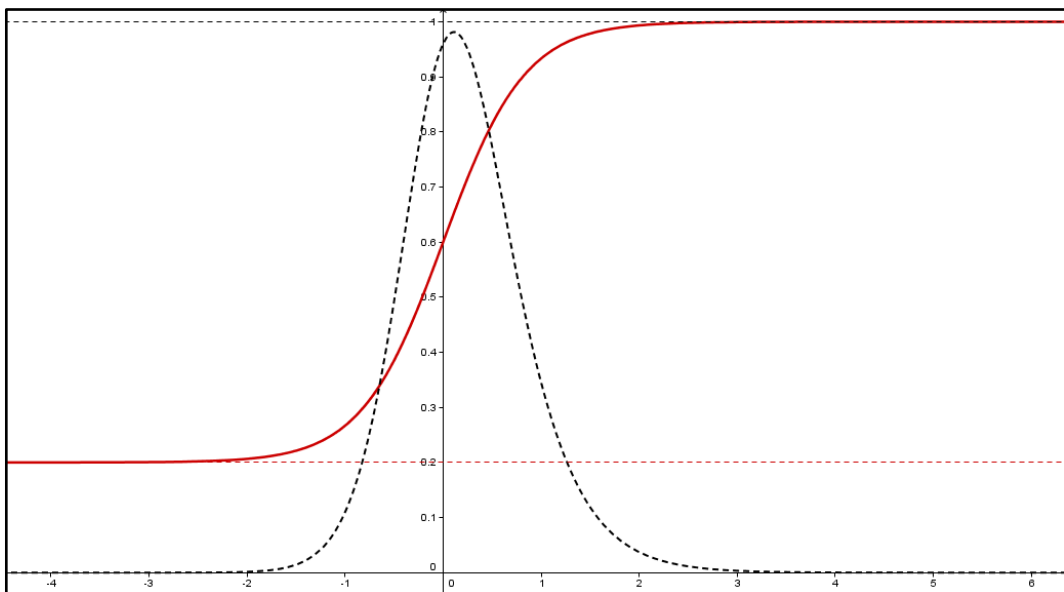


- **VARIANDO O PARÂMETRO  $a$  ( $b$  e  $c$  fixos)**

Nas Figuras 8 e 9, observam-se dois gráficos em que foram fixados os parâmetros  $b$  e  $c$  ( $b = 0$  e  $c = 0,2$ ) e variou-se o parâmetro  $a$  (Figura 8,  $a = 1,5$ ; Figura 9,  $a = 2,4$ ). Nota-se aí uma diferença na curva de informação do item (linha pontilhada) que apresentam uma maior concentração em torno da habilidade média, à medida que o valor do parâmetro  $a$  aumenta. Ou seja, itens com alta discriminação fornecem mais informação na região próxima ao ponto de inflexão da CCI.



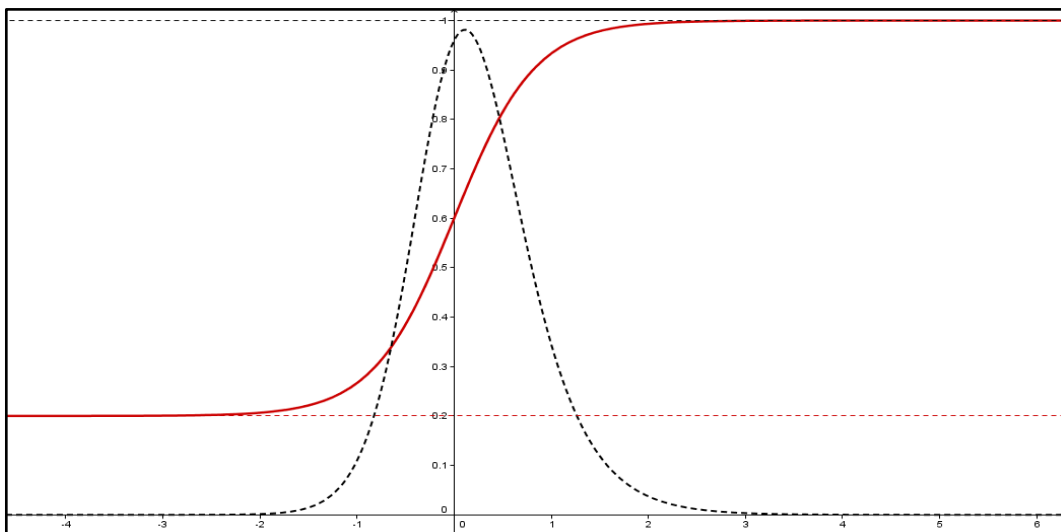
**Figura 8** - CCI e CII construídas juntas, fixando-se os parâmetros  $b$  e  $c$ . O parâmetro  $a$  foi definido em 1,5 ( $a = 1,5$ ).



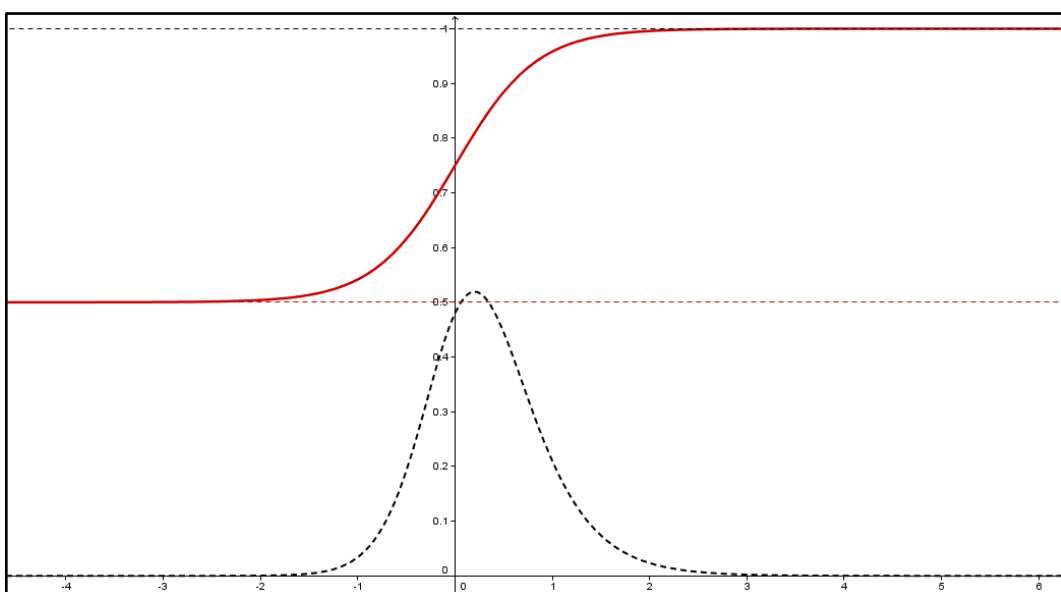
**Figura 9** - CCI e CII construídas juntas, fixando-se os parâmetros  $b$  e  $c$ . O valor do parâmetro  $a$  foi aumentado para 2,4 ( $a = 2,4$ ).

- **VARIANDO O PARÂMETRO  $c$  ( $b$  e  $a$  fixos)**

Nas Figuras 10 e 11, observam-se dois gráficos em que foram fixados os parâmetros  $a$  e  $b$  ( $a = 2,4$  e  $b = 0$ ) e variou-se o parâmetro  $c$  (Figura 10,  $c = 0,2$ ; Figura 11,  $c = 0,5$ ). Verifica-se que o valor máximo da CII diminui à medida que o valor do parâmetro  $c$  aumenta, ou seja, quanto menor o valor de  $c$  (ou seja, menor probabilidade de acerto casual), maior a informação observada nesse item. Desta maneira, itens com baixa probabilidade de acerto casual fornecem mais informações na região próxima ao ponto crítico da CCI.



**Figura 10** - CCI e CII construídas juntas, fixando-se os parâmetros  $a$  e  $b$ . O parâmetro  $c$  foi definido em 0,2 ( $c = 0,2$ ).



**Figura 11** - CCI e CII construídas juntas, fixando-se os parâmetros  $a$  e  $b$ . O valor do parâmetro  $c$  foi aumentado para 0,5 ( $c = 0,5$ ).

Para que se tenha a informação trazida pelo teste completo, visto que ocorre a independência entre os itens, é necessário o somatório de todas as funções de informação do item. Dessa forma, a Função de Informação do Teste é dada por:

$$I(\theta) = \sum_{i=1}^I I_i(\theta)$$

A função de informação do teste também possui uma representação gráfica, chamada Curva de Informação do Teste (CIT), que traz um panorama geral sobre ele, considerando o conjunto de todos os itens e verificando se este teste consegue a melhor avaliação possível de um determinado traço latente.

Todas as análises feitas anteriormente têm como objetivo principal, estudar as habilidades de indivíduos que respondem testes formados por itens construídos a partir do modelo da TRI. Essas habilidades serão expressas numericamente, assumindo qualquer valor do Conjunto dos Números Reais  $\theta \in \mathbb{R}$ , diferente do que ocorre nos testes tradicionais, onde a habilidade do respondente está relacionada com o número de itens que ele acerta, variando então de 0 a n, em que n é o número total de itens. Por essa razão é necessário definir uma escala de valores para essa variável, estabelecendo um ponto central (origem) e qual unidade de medida será adotada para representá-la. Para esse tipo de escala adota-se para origem o valor médio das habilidades dos respondentes e para unidade de medida o desvio-padrão.

### 2.3 CALIBRAÇÃO

A probabilidade de o sujeito analisado fornecer a resposta esperada de certo item depende de sua habilidade e dos parâmetros do item (a probabilidade é uma função da habilidade, parametrizada por características do item) (ANDRADE, TAVARES e VALE, 2000). No entanto, esses parâmetros são geralmente desconhecidos, assim como a habilidade do respondente, sendo necessário definir estes valores a partir das respostas encontradas nos itens. Ao estabelecer a escala de medidas das habilidades, os valores dos parâmetros não sofrerão alteração, mesmo que os itens sejam aplicados a outro grupo respondente, desde que a habilidade desse grupo esteja na mesma escala de medida adotada.

Este processo de estimação de parâmetros de itens é conhecida na TRI como

calibração e pode ser desenvolvida a partir de três perspectivas:

- Conhecem-se os parâmetros dos itens, sendo necessária somente a estimação da habilidade do respondente;
- Conhece-se a habilidade do respondente, necessitando então da estimação dos parâmetros do item;
- Não se tem nenhuma das informações anteriores e, a partir das respostas dos itens, é preciso estimar os parâmetros e as habilidades ao mesmo tempo.

Dentre as três situações apresentadas, a terceira é a mais comum, principalmente quando se trata de aplicar o modelo às avaliações de aprendizagem. Geralmente quando se desenvolve uma avaliação dentro do modelo de TRI tem-se a necessidade de estimar os parâmetros dos itens para então ser possível a estimação das habilidades dos respondentes. Para efetuar esta estimação podem-se utilizar dois métodos: estimar parâmetros e habilidades juntos, na chamada estimação conjunta ou, estimação em duas etapas, em que serão estimados inicialmente os parâmetros dos itens e depois as habilidades dos respondentes.

Conforme Andrade, Tavares e Vale (2000), o diferencial de se trabalhar a estimação conjunta está relacionada com a escala na qual os parâmetros serão estimados. Quando se trabalha a situação em que as habilidades são conhecidas, os parâmetros serão estimados na mesma escala que elas; se apenas os parâmetros são conhecidos, as habilidades serão estimadas na mesma escala que eles. Entretanto, se todos são desconhecidos, há a necessidade de se definir uma escala para efetuar as análises, ou pode ocorrer um problema que os autores chamam de “falta de identificabilidade do modelo”. Isso pode ocorrer nos modelos de três parâmetros (ML3), pois conjuntos de parâmetros diferentes podem gerar o mesmo valor no modelo. Esse problema está relacionado às características do grupo estudado e pode ser eliminada de várias maneiras, como por exemplo, fixando valores para as habilidades.

Enquanto não se define uma escala para as habilidades, não é possível fazer uma gradação entre valores de habilidade encontrados (alta habilidade, baixa, etc.) ou mesmo fazer comparação de valores, verificando se duas habilidades estão próximas ou não. Uma opção para definir essa escala é utilizar a média e o desvio-padrão das habilidades que, conseqüentemente, definirá também a escala dos parâmetros dos itens. Assim é possível eliminar o problema da não-identificabilidade, facilitando o trabalho com as variáveis latentes.

Na maioria das situações que trabalham com TRI tem-se usado o modelo de escala (0,1), em que a média é 0 e o desvio-padrão 1. Através dessa escolha de escala, o parâmetro  $b$  acaba tendo seus valores no intervalo entre  $-2$  e  $2$  ( $b \in [-2,2]$ ) e o parâmetro  $a$  tem seus

valores no intervalo 0 e 2 ( $a \in [0,2]$ ). Apesar disso, vale lembrar que a escolha da escala não interfere na estimação da habilidade do respondente, uma vez que o importante são as relações de ordem existente entre os pontos e, portanto, a probabilidade do respondente fornecer a resposta esperada independe da escala usada. Ainda assim, não será possível fazer qualquer análise das habilidades sem conhecer a escala no qual os parâmetros  $a$  e  $b$  foram estimados.

O processo de estimação pode ser feito de várias maneiras, como vemos em Andrade, Tavares & Vale (2000), sendo necessária a aplicação de processos iterativos, geralmente com o auxílio de softwares estatísticos. Dentre os métodos de calibração, os mais utilizados são o Método da Máxima Verossimilhança e a Inferência Bayesiana, que serão escolhidos conforme a situação que está sendo analisada.

Além da calibração, outro processo que pode ser necessário ao se trabalhar com o modelo de TRI para avaliações é chamado **Equalização**. Este processo é utilizado quando se trabalha com testes diferentes, mas que possuem alguns itens comuns, sendo aplicados a grupos respondentes que deverão ter seus resultados comparados. Dessa forma, as medidas de habilidade encontradas poderão ser comparadas em uma mesma escala de medida.

Após a escolha do método que será utilizado para a calibração e equalização, define-se qual o software será utilizado para realizar as análises, estimar parâmetros, construir as curvas e por fim, estimar as habilidades. Dentre as várias opções de softwares possíveis, optou-se pelo uso do R.

R é o nome dado a um ambiente de desenvolvimento integrado – um programa de computador que reúne características e ferramentas de apoio ao desenvolvimento de software com o objetivo de agilizar esse processo – é também uma linguagem, utilizada para desenvolver cálculos e gráficos. Começou a ser desenvolvido no Departamento de Estatística da Universidade de Auckland na Nova Zelândia, a partir de 1991, pelos cientistas Ross Ihaka e Robert Gentleman, conhecidos por “R&R” (apelido que veio a originar o nome do programa).

O objetivo dos cientistas era criar um programa para usar em suas aulas de laboratórios, usando a linguagem S, utilizada pelo software comercial S-Plus, criado por John M. Chambers da AT&T. Acredita-se que sua distribuição iniciou no ano de 1993, quando algumas de suas cópias foram disponibilizadas através de um sistema de distribuição de softwares estatísticos chamado StatLib.

Um grande diferencial desse software é o fato de ele ser construído de modo colaborativo, em que vários pacotes de funções são inseridos por usuários, sempre ampliando

as possibilidades do programa. Ele realiza uma série de técnicas estatísticas e gráficas e, como utiliza linguagem de programação, permite que cada usuário seja capaz de desenvolver as funções que deseja utilizar, conforme a necessidade. Por ser um software livre, vem sendo cada vez mais utilizado por pesquisadores, além de ser compatível com a grande maioria dos sistemas operacionais existentes.

### 3 MATERIAIS E MÉTODOS

A TRI possui aplicações nas mais variadas áreas, sendo utilizada para estimar diferentes tipos de habilidades. Nesta pesquisa, além da discussão teórica, foi proposta a criação de uma avaliação utilizando este modelo, onde seria feita a estimação de parâmetros e habilidades, procurando ilustrar toda a teoria apresentada anteriormente e, levantando a discussão sobre a aplicação deste modelo no ambiente escolar. Assim, fica claro que o foco da pesquisa não está somente na análise de dados obtidos a partir da aplicação de um teste, mas em todo o desenvolvimento e discussões necessárias para a realização desta ação, além de uma reflexão sobre os resultados encontrados e as vantagens, ou desvantagens, da utilização deste modelo para avaliação de aprendizagem.

Fundamental para o início do trabalho foram as leituras relacionadas ao tema, que mostraram a origem, o desenvolvimento e as aplicações teóricas envolvendo a TRI. A partir dessas leituras foi possível o desenvolvimento de todo o referencial teórico, apresentado na parte inicial do trabalho, e definir quais seriam os objetivos nos quais estaria focado. Além disso, algumas leituras relacionadas com a área de avaliação contribuíram para mostrar o que seria interessante alterar em relação aos modelos já adotados, procurando dessa forma, propor mudanças substanciais e que proporcionem diferenças importantes dentro da área.

Uma vez definido os objetivos da pesquisa, havia a necessidade de se pensar em como seria a ação prática para trabalhar com TRI. Nesse momento há surgimento de algumas questões fundamentais e que serão norteadoras das ações que devem ser tomadas para a realização da pesquisa. Algumas das dúvidas que surgem são: Em qual área aplicar a TRI? Como aplicar a TRI? Qual instrumento será utilizado? Como fazer esse instrumento? Como calibrar os itens? Onde aplicá-lo? Como analisar os resultados encontrados?

Essas questões são fundamentais para a organização da pesquisa e posteriormente, a sua conclusão, visto que indicam qual a direção será tomada e de que maneira isso ocorrerá. O ponto inicial é a definição de onde o modelo de TRI seria aplicado e dessa forma, organizar todos os instrumentos e metodologias necessárias para essa ação. Assim, optou-se por trabalhar com a área da educação, utilizando a TRI para a análise de habilidades em Matemática, através da aplicação de um teste para alguns grupos de alunos. Esta opção vai de encontro com a tendência que se tem observado de utilizar a TRI como avaliação de aprendizagem, buscando uma avaliação mais segura, preocupada em analisar as habilidades específicas dos respondentes, através de itens específicos.

### 3.1 MATERIAIS

Para a criação de itens buscou-se trabalhar com uma única área da Matemática, procurando dessa forma focar num tipo de habilidade e facilitar o processo de aplicação e resolução por parte dos respondentes. A área escolhida foi o “Tratamento da Informação”, tomando-se por base a matriz de competências do Exame Nacional do Ensino Médio (Enem), que leva em consideração os Parâmetros Curriculares Nacionais (PCN). Acredita-se também que ao adotar esse tema, os alunos estariam mais familiarizados, por ser um assunto muito debatido atualmente, principalmente por conta de seu caráter interdisciplinar.

Para a elaboração de uma avaliação no modelo da TRI é necessário um grande cuidado com a escolha dos itens, que deverão estar dentro de certos padrões, como possuir diversos graus de dificuldade, a mesma quantidade de alternativas e, cada item deve avaliar uma única habilidade (garantindo a unidimensionalidade do modelo). Os parâmetros dos itens devem estar bem definidos e isso ocorre através do processo de calibração, que será feito a partir da aplicação de um pré-teste a uma amostra de respondentes. Esse pré-teste tem por objetivo verificar se os itens escolhidos estão dentro dos parâmetros necessários para serem utilizados num modelo de TRI, evitando incertezas nas análises ou, possíveis anulações de itens após aplicação da avaliação. A partir da análise dos resultados deste pré-teste serão definidas mais seguramente o grau de dificuldade e a discriminação dos itens e então, serão escolhidos aqueles que farão parte da avaliação.

Definido o tema do teste, o grupo respondente também começava a ser delimitado, uma vez que houve a preferência de se aproximar do modelo de avaliação proposto pelo Enem (modelo que utiliza a TRI). Inicialmente pensou-se em aplicar o teste somente aos alunos do Ensino Médio de escolas públicas e particulares do município de Rio Branco, mas devido ao período de aplicação ter coincido com as férias escolares, foi necessária uma readequação do grupo respondente, de forma que fosse possível conseguir uma amostra maior. Dessa forma, para responder o teste foram escolhidos cinco grandes grupos formados pelos seguintes alunos:

- **1º grupo:** alunos que cursam o terceiro ano do Ensino Médio no Colégio de Aplicação da Universidade Federal do Acre (CAp – UFAC) e estão se preparando para fazer o Enem e tentar ingresso no nível superior;
- **2º grupo:** alunos que cursam o quarto ano do Ensino Médio Integrado ao Técnico no Instituto Federal do Acre (IFAC), com objetivos semelhantes aos do



grupo anterior;

- **3º grupo:** alunos que cursam o terceiro ano de Ensino Médio na Escola Estadual Professora Clícia Gadelha
- **4º grupo:** acadêmicos no primeiro período de cursos superiores no IFAC, que ingressaram através do Enem (calouros);
- **5º grupo:** acadêmicos de diferentes períodos e cursos dentro da UFAC.

A escolha dos grupos levou em consideração a crença de que estariam mais familiarizados com o conteúdo proposto no teste e dessa forma, suas habilidades em Matemática poderiam ser quantificadas de forma mais segura. Essa crença baseia-se no fato de que os respondentes dos 1º, 2º e 3º grupos estariam estudando esse assunto na escola, ou se preparando para a realização do Enem. No caso dos calouros, como acabaram de fazer Enem, não teriam dificuldades, visto que estavam em ritmo de estudos para o ingresso no nível superior. O último grupo foi escolhido porque é composto por alunos que cursam diferentes graduações, mas no momento cursam disciplinas que contemplam os conteúdos propostos no teste. Aplicar o teste para respondentes que não estivessem familiarizados ao conteúdo poderia gerar resultados que não seriam interessantes à pesquisa, pois, como havia pouco tempo para aplicação, os respondentes não teriam a oportunidade de se preparar previamente (estudando o tema) e a probabilidade de fornecerem a resposta correta seria muito baixa. Cabe ainda salientar que a escolha desses grupos também levou em consideração a acessibilidade a eles, uma vez que estão vinculados a instituições que permitiriam o desenvolvimento da pesquisa, pois os pesquisadores possuem vínculos com estas.

Na criação do pré-teste optou-se por utilizar itens prontos, recolhidos através de consulta a sites de internet e outros materiais didáticos. Essa escolha teve a intenção de aproveitar melhor o tempo, direcionando-o para a aplicação e análises dos resultados, ao invés de gastá-lo na formulação dos itens, uma vez que se pretendia aplicar o pré-teste entre alunos de escolas públicas antes do início das férias escolares. Para a escolha dos itens deu-se preferência para questões aplicadas em diversos vestibulares, de instituições públicas ou privadas, provas do Enem, provas de concursos públicos e Olimpíada Brasileira de Matemática (OBMEP), tendo a preocupação de formatá-las quando necessário, para que tivessem as características de um item de TRI, principalmente no que se refere a avaliação de uma única habilidade a cada item proposto. É importante salientar que a elaboração dos itens é um processo que requer mais tempo, pois algumas características devem ser observadas uma delas é que o elaborador já deve começar a pensar nas características da TRI desde o momento inicial da sua avaliação.

O teste foi composto por 18 (dezoito) itens objetivos, com cinco alternativas cada, em que havia somente uma correta (garantindo a dicotomia dos itens). Esses itens foram separados em três assuntos, dentro do tema escolhido, sendo seis itens sobre Probabilidade, seis sobre Análise Combinatória e seis sobre Estatística, havendo também variação no nível de dificuldade dos itens, segundo a opinião inicial dos pesquisadores, sem a utilização da TRI ainda. A dificuldade foi classificada numa escala de 1 a 5, onde o valor 1 foi atribuído às questões consideradas bem fáceis, de resolução direta; 2 para as questões fáceis mas de resolução não tão direta; 3 para questões medianas; 4 para questões um pouco mais difíceis e 5 para questões consideradas difíceis. O pré-teste continha duas questões classificadas como medianas (nível 3) e uma questão de cada um dos outros níveis, para cada um dos assuntos, sendo que a ideia dessa classificação foi elaborada a partir de Professor Cardy (2014) um dos sites de onde os itens foram coletados. Para classificar a dificuldade dos itens seguiram-se dois caminhos:

- Os itens retirados de Professor Cardy (2014) já vinham com a classificação indicativa do grau de dificuldade. A classificação lá encontrada era feita a partir da opinião dos usuários do site, que votavam no grau de dificuldade que acreditavam que melhor se encaixava no item. A partir de então se resolvia o item para confirmar essa classificação ou propor uma nova;
- Os itens retirados de outros sites, em que não havia classificação indicativa, eram resolvidos e discutidos para chegar a uma categorização, procurando analisar conforme os modelos vistos em Professor Cardy (2014), para que não houvesse muita discrepância entre os mesmos.

Para a aplicação do pré-teste em cada grupo foi disponibilizado um tempo de uma hora e meia (90 minutos), sendo os respondentes orientados a não deixarem itens em branco. Caso não soubessem respondê-los, deveriam marcar qualquer alternativa, uma vez que o acerto casual está previsto na análise dos resultados. A aplicação do teste às turmas ocorreu de forma tranquila e todos os respondentes se disponibilizaram a participar da pesquisa, entendendo a importância de suas respostas. Mesmos alunos que não estavam muito interessados em participar responderam ao teste, contribuindo para a análise do acerto casual que será feita a partir dos resultados.

Com os testes respondidos em mãos inicia-se a fase de análise das respostas encontradas, verificando se o teste cumpriu seu objetivo de selecionar e calibrar itens para a elaboração de um modelo de avaliação. A primeira coisa a ser feita é verificar se todos os itens foram compreendidos conforme se esperava, ou se havia algum problema de clareza nos

enunciados. Caso algum dos itens não tenha sido compreendido, ou tenha apresentado algum problema em seu enunciado, ou alternativas, será descartado da avaliação, por se acreditar que este não é um bom item. Essa análise começa durante a aplicação do pré-teste, questionando os alunos se há alguma dúvida nos enunciados das questões, logo que o teste é entregue a eles, ou ainda, observando algumas perguntas que são feitas durante o teste, que podem demonstrar esse tipo de dúvida.

Na primeira turma em que o pré-teste foi aplicado, um dos respondentes questionou o enunciado do item 12 (Figura 12), pois não havia nenhuma especificação se os números utilizados no item deveriam ser diferentes de 1 e, caso essa condição não fosse garantida, não haveria resposta correta entre as alternativas.

(FGV 2008 nível 5) Sendo  $x$ ,  $y$  e  $z$  três números naturais tais que  $x \cdot y \cdot z = 2310$ , o número de conjuntos  $\{x, y, z\}$  diferentes é

a) 32                      b) 36                      c) 40                      d) 43                      e) 45

**Figura 12** - Item 12 do pré-teste

Após análise do questionamento e verificação do gabarito da questão, achou-se interessante inserir a informação, visto que, a resposta creditada como correta no gabarito dependia dessa condição. Em todas as outras aplicações essa informação foi inserida no item, para não acarretar dúvidas entre os respondentes.

Em uma das turmas de aplicação foi solicitado que se fizesse a correção do pré-teste junto aos respondentes, com intenção de esclarecimento de dúvidas e debates sobre o tema. Essa ação foi muito interessante, pois foi possível verificar como os alunos classificaram cada item quanto ao grau de dificuldade encontrado por eles e comparar com a classificação inicialmente proposta pelos pesquisadores. Além disso, constatou-se a partir de dúvidas levantadas pelos respondentes que, a questão de número 3 (Figura 13), envolvendo o tema probabilidade, não possuía resposta correta. Após discussão, observou-se que não era possível alterar a questão para que tivesse resultado, sendo necessária a anulação da mesma e sua exclusão da análise.

(CESGRANRIO 2008 – Adaptada) Em uma urna há 5 bolas verdes, numeradas de 1 a 5, e 6 bolas brancas, numeradas de 1 a 6. Dessa urna retiram-se, sucessivamente e sem reposição, duas bolas. Qual a probabilidade de, na primeira extração a bola sacada ser verde e na segunda ser um número par?

a)  $\frac{2}{11}$                       b)  $\frac{3}{11}$                       c)  $\frac{5}{22}$                       d)  $\frac{25}{121}$                       e)  $\frac{3}{22}$

**Figura 13** - Item 3 do pré-teste.

Para otimizar o trabalho com os resultados encontrados de modo que não houvesse confusão com as informações, achou-se mais prudente renumerar os itens do teste após a exclusão do item 3, considerando a partir de então, que o teste possuía 17 questões, ao invés de 18. Com isso a classificação do grau de dificuldade das questões, antes da análise com a TRI passa a ter a seguinte organização, apresentada na Tabela 1.

**Tabela 1** - Classificação dos itens em relação à dificuldade, antes da análise com TRI.

GRAU DE DIFICULDADE	QUESTÕES
1	1, 6,12
2	2, 7, 14
3	3, 8, 9, 13,15
4	4, 10, 16
5	5, 11, 17

Em conversa com os respondentes após a aplicação, notou-se que a maioria encontrou dificuldades nas questões que foram classificadas na escala como 4 ou 5 (questões que eram para ser mais difíceis), principalmente nas que foram classificadas como 5. Aparentemente, isso mostra que as questões classificadas como difíceis foram as que realmente exigiram maiores habilidades dos respondentes. Todavia, só a partir da análise das respostas de todos os pré-testes é que será possível fazer uma estimativa mais precisa do parâmetro dificuldade.

Após esse momento ocorreu a correção dos itens, utilizando um gabarito previamente elaborado e a atribuição de um número a cada um dos testes, como forma de identificação, sem relacionar este número aos nomes dos respondentes, verificando-se que foram 216 respondentes. O passo seguinte é a representação dessas respostas na forma de uma tabela (matriz) de valores binários (0 para respostas erradas e 1 para respostas certas) que é utilizada para a estimação dos parâmetros e habilidades, sendo cada linha dessa matriz um vetor associado a cada um dos respondentes que participaram da pesquisa. Essa matriz é composta por 216 linhas e 17 colunas e será apresentada conforme o Apêndice B.

### 3.2 METODOLOGIA

Como visto anteriormente, optou-se por utilizar o software R para fazer as análises estatísticas e estimação de parâmetros desta pesquisa. A escolha do programa levou em consideração dois fatores principais: o fato de ser um software livre e possuir pacotes de funções que já são aplicadas em análises envolvendo a TRI. Os pacotes de funções utilizados foram o *ltm*, o *irtoys* e o *CTT*, efetuando diferentes análises. Foi adotado como método de estimação o Método da Máxima Verossimilhança Marginal, através do R, que fez todas as estimações, usando métodos iterativos.

Entretanto, para a realização da parte gráfica da pesquisa optou-se pela utilização do software GeoGebra, uma vez que este já vinha sendo usado desde o início da pesquisa para a construção das curvas, e por isso seus comandos já eram mais conhecidos. Além disso, ele apresenta uma melhor visualização das curvas facilitando as análises, além da possibilidade de combinar a construção de várias curvas num mesmo eixo. Por fim, o fato de possuir comandos dinâmicos facilitam a visualização das mudanças que ocorrem nas curvas quando alteramos seus parâmetros, sendo possível verificar o formato das curvas para itens considerados ideais ou mesmo impróprios.

Ainda para auxiliar nas análises dos parâmetros fornecidos pelo R, foi criada uma escala de dificuldades, tomando como referência a distribuição logística com parâmetros de dificuldade igual a zero e discriminação 1 (logística padrão). O intuito seria classificar os itens em relação ao seu grau de dificuldade (parâmetro  $b$ ) de forma clara e objetiva, facilitando as análises e proporcionando posteriormente, uma forma de organizá-los, usando este parâmetro como um critério. Neste sentido foram considerados os quantis a 20%, 40%, 60% e 80% para fazer a classificação. Na Tabela 2 seguem os resultados que serão usados no decorrer do trabalho.

**Tabela 2** - Escala de dificuldades utilizada como referência no trabalho.

Valor do parâmetro $b$ (dificuldade)	Grau de dificuldade
$b < -1,38$	Muito fácil
$-1,38 \leq b < -0,41$	Fácil
$-0,41 \leq b < 0,41$	Intermediário
$0,41 \leq b < 1,38$	Difícil
$b \geq 1,38$	Muito difícil

## 4 RESULTADOS E DISCUSSÃO

Após definir quais os materiais e qual a metodologia seria usada para a realização do trabalho, inicia-se a fase das análises e discussões, utilizando os resultados encontrados durante a aplicação dos testes aos grupos e os softwares anteriormente citados (R e GeoGebra). As análises serão feitas na perspectiva de definir quais seriam os melhores itens para a composição de uma avaliação, a partir das características ideais vistas no Referencial Teórico, utilizando também a Tabela 2 para classificar quanto ao grau de dificuldade.

### 4.1 ESTIMAÇÃO DOS PARÂMETROS A PARTIR DO R: CALIBRAÇÃO

Iniciaram-se as análises com o R utilizando o pacote *ltm*, que trabalha com análises relacionadas com a Teoria Clássica dos Testes (TCT). Apesar de a pesquisa estar focada na metodologia da TRI, algumas informações interessantes podem ser obtidas a partir da TCT, reforçando a ideia que de uma metodologia não tem a intenção de superar a outra, mas apenas oferecer novas abordagens e possibilidades de análises diferenciadas, buscando melhorar situações e análises que a TCT não conseguia fazer anteriormente. Através do pacote *ltm* foi construída a tabela de frequência dos escores brutos dos testes, associando o número de acertos à quantidade de respondentes que atingiram aquele valor. A Tabela 3 a seguir traz essas informações, mostrando que houve uma grande concentração de respondentes abaixo dos 8 acertos, que seria a metade do teste.

**Tabela 3** - Frequência absoluta dos acertos.

ACERTOS	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
FREQ.	0	3	12	16	25	22	33	34	38	21	4	6	2	0	0	0	0	0

Outra análise feita pelo pacote *ltm* foi o percentual de acertos e erros de cada um dos itens do teste, apresentado na Tabela 4.

**Tabela 4** - Tabela com percentual de acertos e erros calculados pelo R.

ITENS	ERROS (%)	ACERTOS (%)
1	65,28	34,72
2	34,72	65,28
3	70,37	29,63
4	73,61	26,39
5	77,31	22,69
6	60,65	39,35
7	74,54	25,46
8	62,04	37,96
9	66,2	33,8
10	71,3	28,7
11	80,09	19,91
12	43,98	56,02
13	77,31	22,69
14	43,06	56,94
15	45,37	54,63
16	50,46	49,54
17	81,48	18,52

A partir dos dados constantes na Tabela 4 observou-se que o percentual de acertos das questões não foi muito alto (o item com maior índice de acerto foi o 2, com 65,28%), enquanto o percentual de erros foi maior (o item 17 apresentou 81,48%) mostrando que os respondentes não tiveram um alto desempenho no pré-teste. Ainda observou-se que as questões 17, 11 e 5 foram aquelas que tiveram o maior índice de erro entre os respondentes, enquanto as questões 2, 14 e 12 foram às que tiveram o maior índice de acerto. Comparando os itens em relação à classificação de dificuldade feita inicialmente, as questões com maior percentual de erro foram àquelas classificadas como mais difíceis (nível 5) enquanto as questões com maior percentual de acerto foram as classificadas como mais fáceis (nível 1 e 2), numa análise típica de quando se trabalha com a teoria clássica.

As tabelas construídas a partir da teoria clássica têm função de auxiliar as análises feitas no modelo da TRI, mostrando o desempenho dos respondentes em relação as quantidades de acertos e erros. Essa comparação pode auxiliar no momento de analisar cada item em relação aos seus parâmetros, que é o foco principal da TRI.

Na sequência utilizou-se o pacote *irt*, em que foi considerado o modelo logístico de três parâmetros (modelo da TRI) com a intenção de fazer a estimação dos três parâmetros de cada item (discriminação, dificuldade e acerto casual). A Tabela 5 apresenta os resultados.

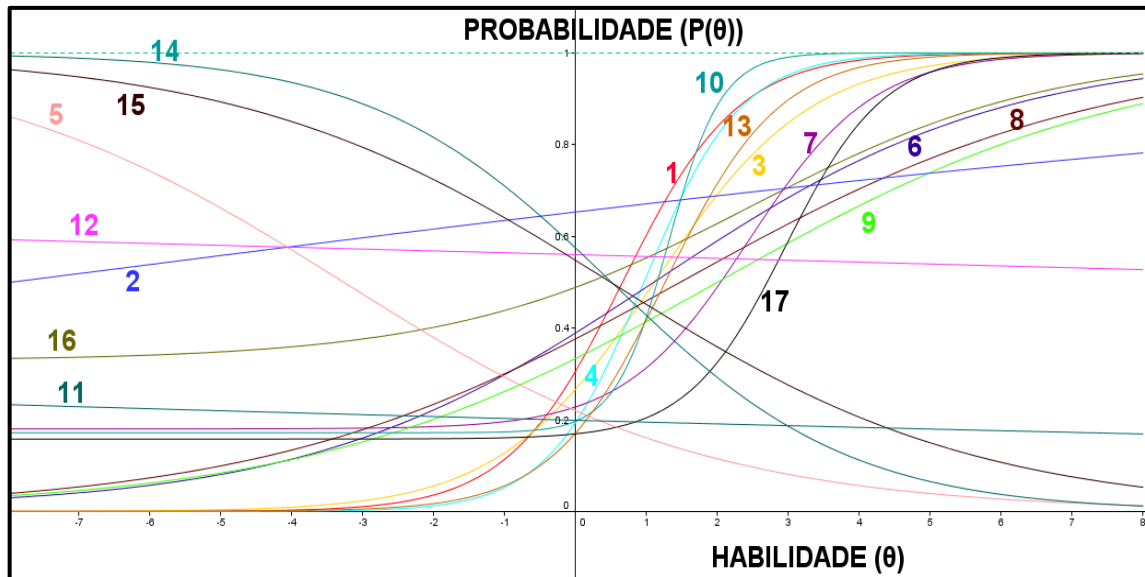
**Tabela 5** - Parâmetros dos itens após análise feita pelo R.

<b>Itens</b>	<b>Discriminação</b>	<b>Dificuldade</b>	<b>Acerto casual</b>
1	1,2356	0,6633	0,0000
2	0,0817	- 7,0610	0,0353
3	0,9095	1,1134	0,0001
4	1,4718	0,9667	0,0000
5	- 0,3877	- 3,2706	0,0003
6	0,4107	1,1410	0,0070
7	1,1406	2,4383	0,1801
8	0,3426	1,4903	0,0021
9	0,3483	2,0311	0,0049
10	2,6598	1,3105	0,1718
11	- 0,0275	-54,8422	0,0216
12	- 0,0174	7,5294	0,0587
13	1,2515	1,2660	0,0008
14	-0,5930	0,5082	0,0002
15	-0,3863	0,4924	0,0009
16	0,4724	2,4615	0,3289
17	1,4481	2,9646	0,1580

#### **4.2 ANÁLISE DOS ITENS ATRAVÉS DE SUAS CURVAS: USANDO GEOGEBRA.**

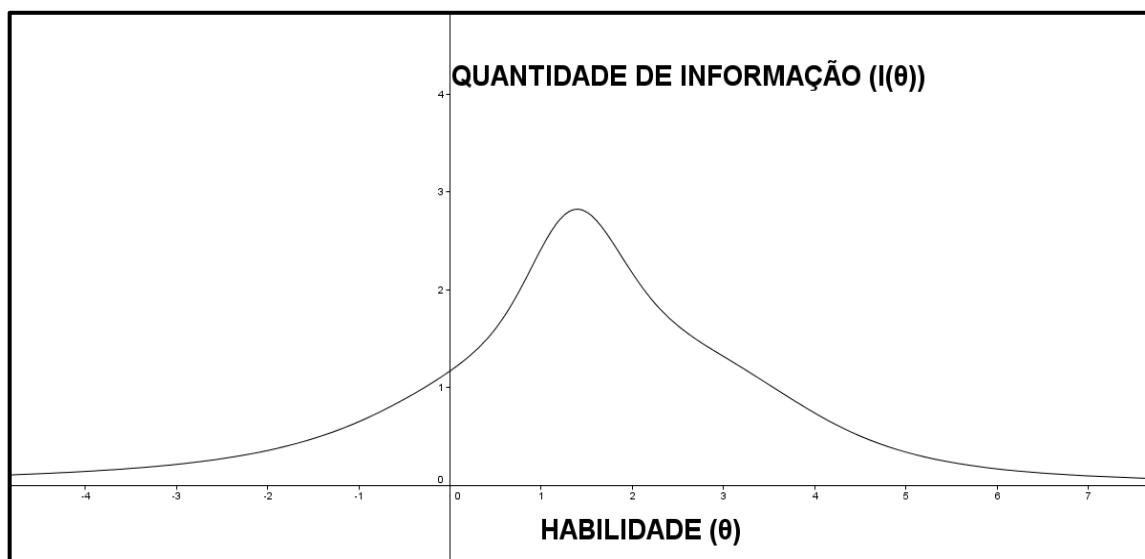
Com as informações da Tabela 5, foi possível fazer a construção das Curvas Características de cada um dos itens do teste (CCI), utilizando o software GeoGebra (Figura 14). Essas curvas foram construídas juntas nos eixos que relacionam habilidade dos respondentes (eixo horizontal) e probabilidade da resposta correta (eixo vertical). Cada curva apresenta o número do item que está representado e foi construída para um intervalo de - 8 a 8 no eixo das habilidades. Para essa construção e todas que se seguiram no GeoGebra, foi necessário fazer o arredondamento dos valores dos parâmetros para duas casas decimais, facilitando a manipulação e a visualização das curvas.





**Figura 14 -** Curvas Características de todos os itens do pré-teste

Além das curvas características de cada item, ainda há a possibilidade de se trabalhar com a função de informação do teste, encontrada a partir do somatório de todas as funções de informação dos itens e, representá-la graficamente (Curva de Informação do Teste), como se pode observar na Figura 15. Essa curva, como já citado anteriormente, pode trazer informações gerais a respeito do teste, mostrando se ele consegue obter muita informação a respeito das habilidades dos respondentes. Através do GeoGebra foi possível a construção dessa curva, utilizando eixos que relacionam a habilidade do respondente (eixo horizontal) e quantidade de informação fornecida pelo teste (eixo vertical).



**Figura 15 -** Curva de Informação do Teste

Usando a escala de dificuldades da Tabela 2, é possível classificar o teste, a partir da curva de informação apresentada na Figura 15, mostrando um teste com dificuldade alta (a curva está mais concentrada próxima ao intervalo  $1 < b < 2$ ) indicando que há pouca variabilidade no grau de dificuldade dos itens. De forma geral, o teste apresenta uma qualidade razoável, sendo possível trabalhar a estimação das habilidades dos respondentes.

Com base nas informações obtidas através da Tabela 5 é possível verificar quais itens presentes no pré-teste podem ser considerados para se realizar uma boa avaliação de conhecimento. A CCI de um item deve ser uma função monotônica crescente, que representa o aumento da probabilidade de acerto do item a medida que se aumenta a habilidade do respondente. Observando a Figura 14, foram encontrados cinco itens cuja CCI apresenta comportamento decrescente, que significaria em teoria que quanto maior a habilidade do respondente, menor a probabilidade dele fornecer a resposta correta. Ao verificar a Tabela 5 com as estimativas dos parâmetros notou-se que esses mesmo itens apresentam o parâmetro discriminação (parâmetro  $a$ ) com valores negativos, inviabilizando a utilização destes itens. Dessa forma, os itens **5, 11, 12, 14 e 15** serão desconsiderados na elaboração de uma possível avaliação, pois não trazem contribuição para a estimação de habilidades dos respondentes.

Continuando a análise, outras características foram verificadas, procurando dessa forma encontrar quais dos itens restantes seriam os ideais para compor uma boa avaliação no modelo da TRI. Uma das possibilidades de análise é através da Função de Informação do Item, que mostra a quantidade de informação que o item apresenta (de forma visual). Esta função, que também utiliza os parâmetros dos itens, gera a curva conhecida como Curva de Informação do Item (CII), que está na mesma escala da CCI e por isso, pode ser construída junto dela. Ao comparar as duas curvas foi possível verificar quais itens do pré-teste não eram interessantes à avaliação, pois não apresentavam muita informação sobre a habilidade do respondente.

Um bom item apresenta muita informação e isso ocorre quando ele apresenta alta discriminação (parâmetro  $a$ ) e baixa probabilidade de acerto casual (parâmetro  $c$ ). Nas Figuras 16 a 27 serão apresentadas as duas curvas construídas em um mesmo gráfico (CCI e CII), referentes a cada um dos itens restantes do pré-teste e na sequência, uma análise de seu comportamento, comparando com a classificação inicial sobre grau de dificuldade (Tabela 1), percentual de acertos e erros (Tabela 4), com as estimativas dos parâmetros dos itens (Tabela 5) e a nova classificação de dificuldade dos itens, baseada no parâmetro  $b$  (Tabela 2). A CCI estará representada pela curva com linha cheia enquanto a CII será representada pela linha tracejada.

### I. Item 1:

O item 1 teve um percentual de acertos de 34,72%, que poderia ser classificado como um percentual intermediário (em relação aos outros percentuais de acertos encontrados). Ele apresenta uma boa discriminação ( $a = 1,24$ ), que se reflete no formato de sua CCI, na Figura 16. Seu parâmetro dificuldade ( $b = 0,66$ ) o classifica como um item difícil, segundo a classificação proposta na Tabela 2, diferente da classificação proposta na Tabela 1, em que seria um item muito fácil. Seu parâmetro de acerto casual foi zero ( $c = 0$ ), que é uma boa característica para um item e sua CII apresenta informação intermediária (o valor máximo desta se aproxima de 0,4). Dessa forma, esse item apresenta boas características e poderia ser utilizado em uma avaliação.

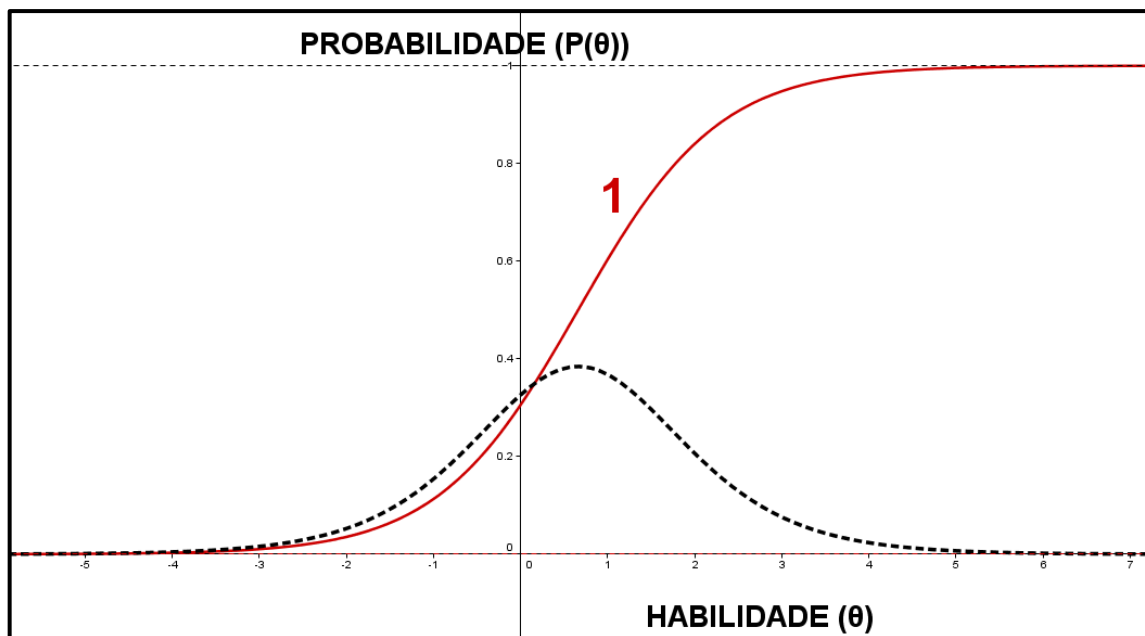
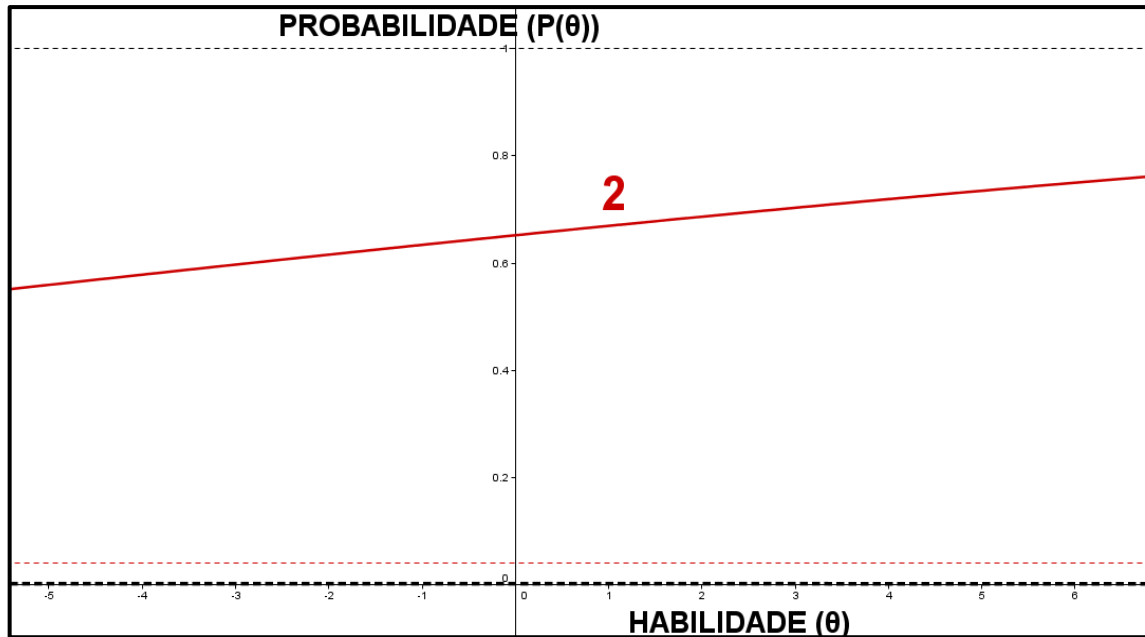


Figura 16 - CCI e CII do item 1 do pré-teste..

### II. Item 2:

O item 2 apresentou a maior quantidade de acertos do pré-teste, com um percentual de 65,28%, levando-se a acreditar, numa análise inicial, que fora considerado o item mais fácil pelos respondentes. Seu parâmetro dificuldade foi o segundo mais baixo de todo o teste ( $b = -7,06$ ) e o primeiro entre os itens que não foram excluídos inicialmente, podendo ser classificado como uma questão muito fácil, diferente da classificação inicial, onde fora classificado como fácil. Entretanto, esse item não nos traz muita informação, como pode-se ver pela sua CII na Figura 17, que apresenta comportamento semelhante a uma reta, abaixo da

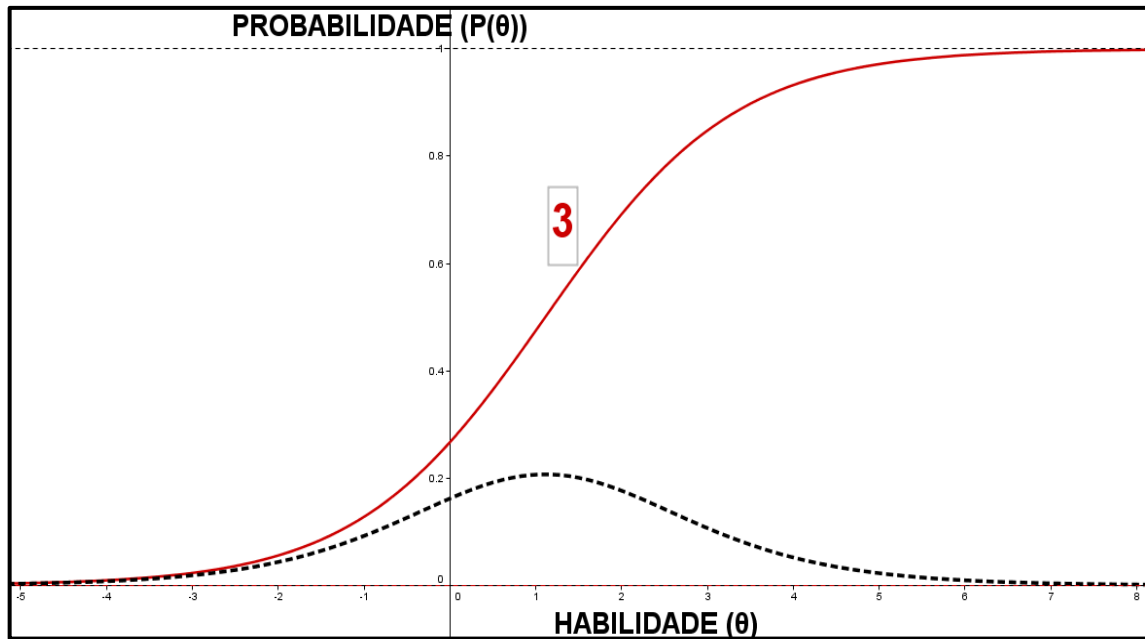
assíntota inferior, que representa o valor do acerto casual, que também foi baixo ( $c = 0,04$ ). Além disso, seu parâmetro de discriminação foi muito baixo ( $a = 0,08$ ), trazendo um item que não gera muita informação, não sendo um bom estimador de habilidades e por isso, será desconsiderado para uma possível avaliação.



**Figura 17** - CCI e CII do item 2 do pré-teste..

### III. Item 3:

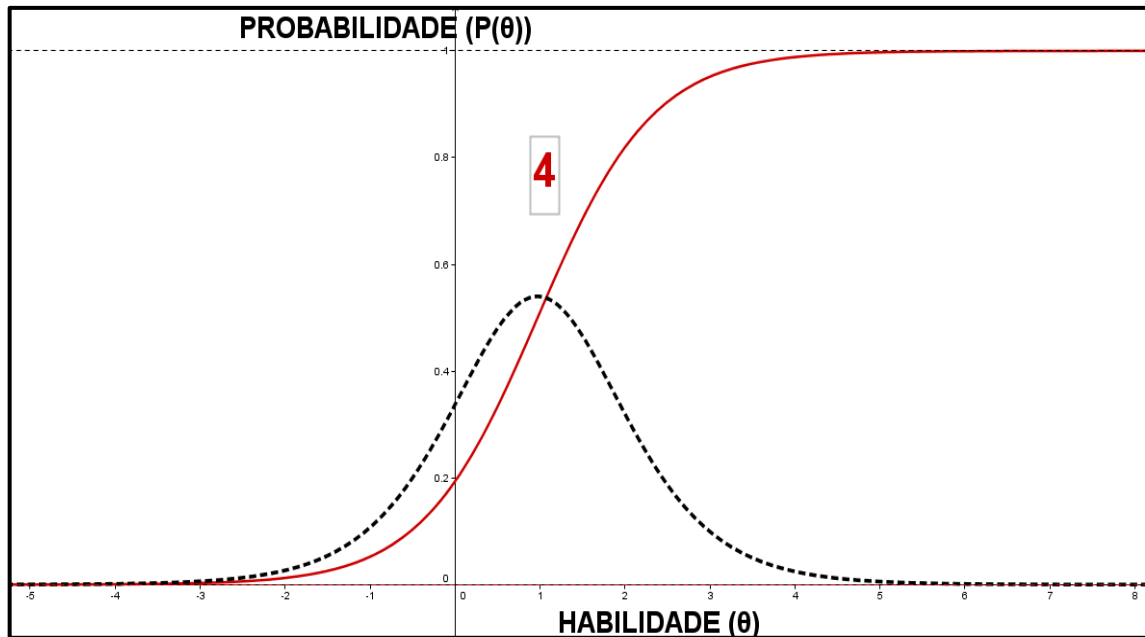
O item 3 apresentou um percentual de acerto de 29,63%, que pode ser considerado um pouco abaixo da média de acertos dos itens, mostrando que os respondentes tiveram um pouco mais de dificuldade para resolvê-lo. Foi classificado inicialmente como um item de dificuldade intermediária mas, após a estimação do seu parâmetro dificuldade ( $b = 1,11$ ) recebe uma nova classificação, seguindo a Tabela 2, sendo considerado um item difícil. Mesmo que sua discriminação não tenha sido muito alta ( $a = 0,91$ ) e sua CII apresentar pouca informação (Figura 18), seu acerto casual foi zero ( $c = 0$ ) e a combinação dessas características tornam esse item interessante para estar em uma avaliação.



**Figura 18** - CCI e CII do item 3 do pré-teste..

#### **IV. Item 4:**

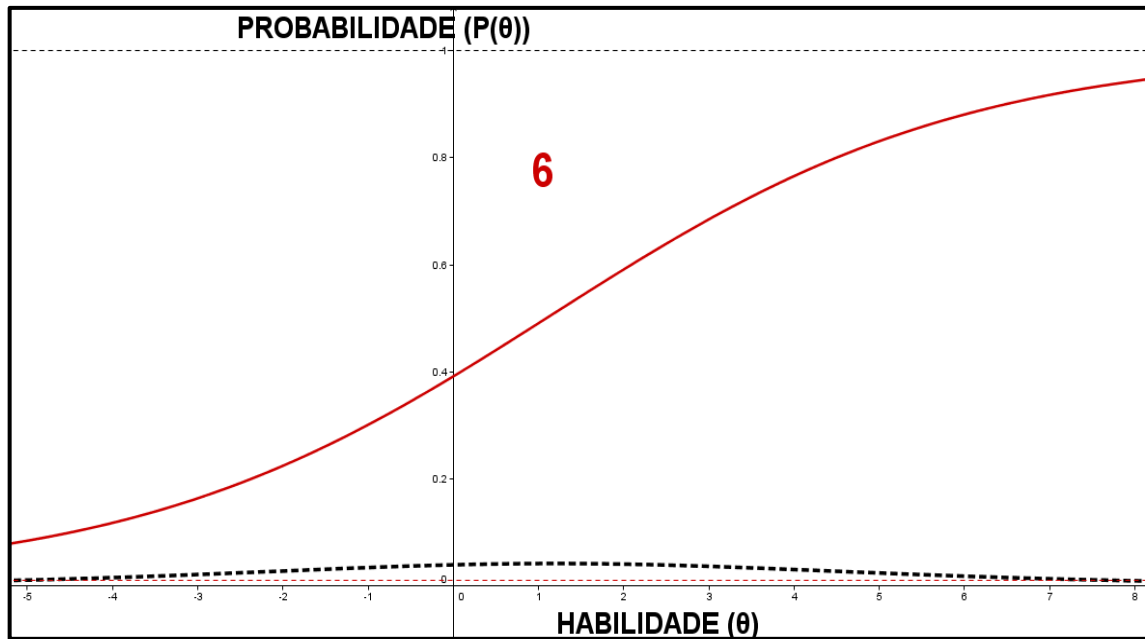
O item 4 apresentou um baixo percentual de acerto (26,39%) e poderia ser considerado, numa análise superficial, um item de dificuldade um pouco acima da intermediária, conforme a classificação inicialmente proposta pelos pesquisadores mas, a estimativa de seu parâmetro de dificuldade ( $b = 0,97$ ) a classifica como um item difícil. Ele apresenta uma alta discriminação ( $a = 1,47$ ), sendo a segunda maior de todo o teste, e isso já fica evidente na sua CCI, mais inclinada, e na sua CII, cujo valor máximo no gráfico ultrapassa a CCI (Figura 19). Seu acerto casual foi estimado em zero ( $c = 0$ ), que é uma boa característica. Logo, esse item possui bons parâmetros e pode ser incluído em uma avaliação.



**Figura 19** - CCI e CII do item 4 do pré-teste.

#### V. Item 6:

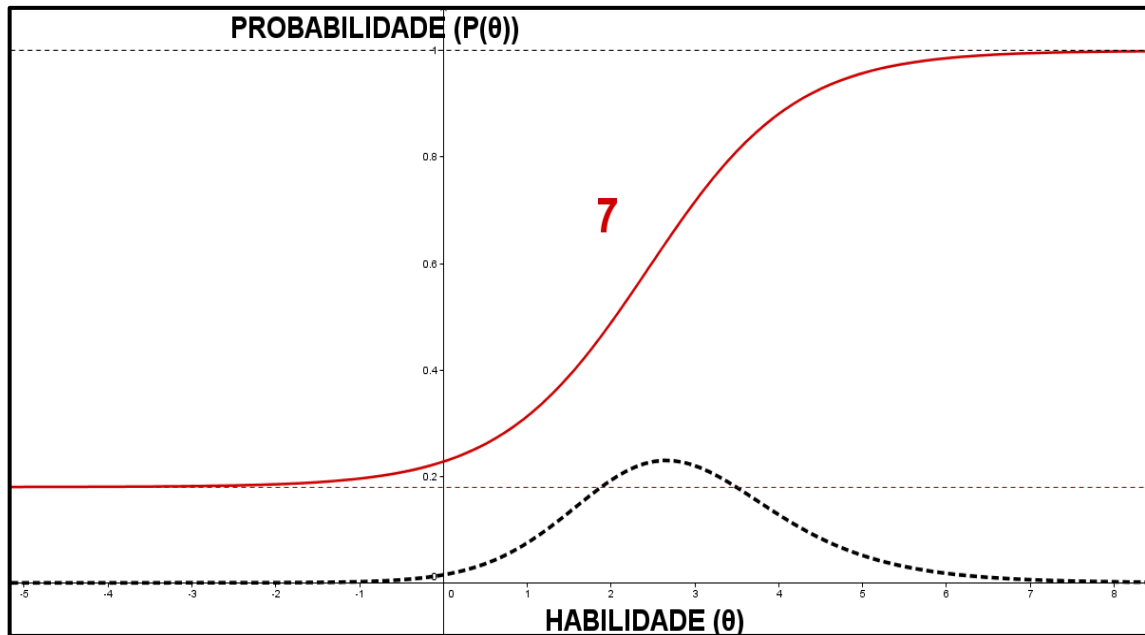
O item 6 apresentou um percentual de acertos de 39,35%, que pode ser considerado um valor numa faixa intermediária de acertos. A estimação de seu parâmetro dificuldade ( $b = 1,14$ ) indica que o item será classificado como difícil, que difere da classificação proposta inicialmente, em que seria um item fácil, de resolução direta. Sua discriminação foi baixa ( $a = 0,41$ ) e isso se reflete em suas curvas (Figura 20), uma CCI com pouca inclinação e uma CII com valor máximo baixíssimo, quase coincidente com o eixo das habilidades, mostrando que o item apresenta pouquíssima informação. Seu parâmetro de acerto casual foi baixo ( $c = 0,01$ ), mas mesmo assim, não torna esse item viável para utilização, sendo aconselhável a sua exclusão.



**Figura 20** - CCI e CII do item 6 do pré-teste.

#### **VI. Item 7:**

O item 7 apresentou 25,46% de acertos, que pode ser considerado um percentual baixo, quando comparado com os dos outros itens, passando uma ideia inicial de ser um item com dificuldade intermediária ou um pouco acima desta. Seu parâmetro de dificuldade foi alto ( $b = 2,44$ ) e o classifica como um item muito difícil, contrariando a primeira classificação, em que foi definido como um item fácil mas, de resolução não direta. Suas CCI e CII confirmam o maior grau de dificuldade, visto que estão mais deslocadas para a direita (Figura 21), se comparadas com as outras curvas. Sua discriminação apresentou um valor um pouco acima do intermediário ( $a = 1,14$ ) quando comparamos com os outros itens encontrados, gerando uma CCI com uma inclinação razoável. Apresentou o parâmetro de acerto casual alto ( $c = 0,18$ ), sendo o segundo maior de todo o teste e isso fez com que sua CII não fosse tão expressiva, mostrando que o item não oferecia muitas informações. Numa análise geral, esse item poderia ser considerado um item mediano, mas que ainda poderia ser utilizado dentro de uma avaliação.

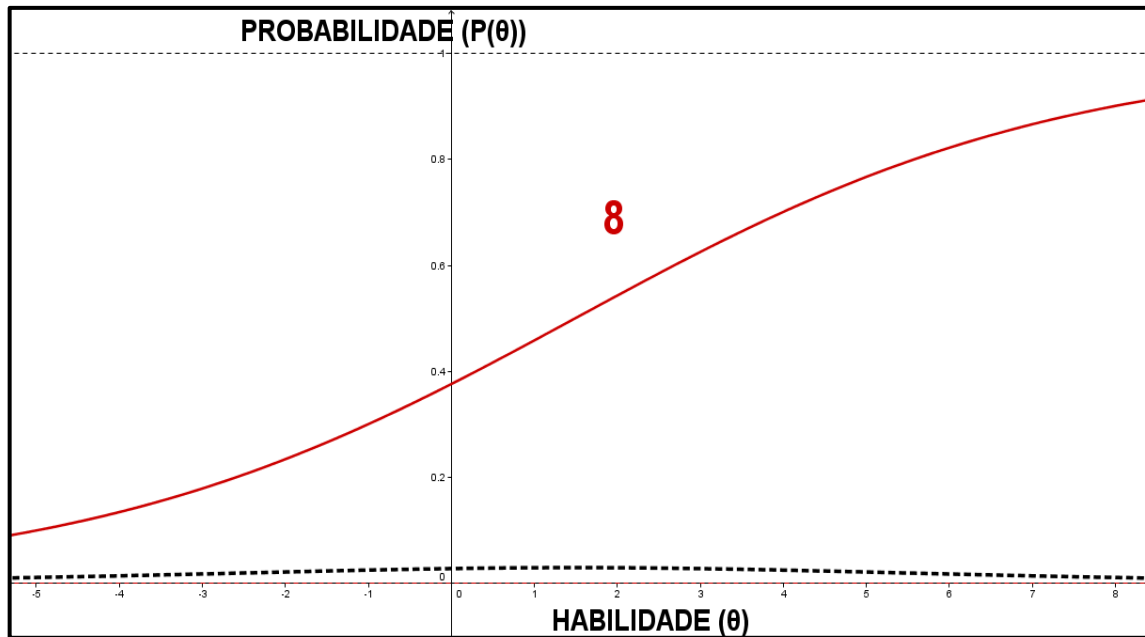


**Figura 21** - CCI e CII do item 7 do pré-teste.

## VII. Item 8

O item 8 apresentou 37,96% de acertos, que ao ser comparado com os outros, pode ser considerado como uma quantidade mediana de acertos e, numa análise superficial, poderia ser classificada como um item de dificuldade intermediária. Seu parâmetro de dificuldade ( $b = 1,49$ ) indica que seria um item muito difícil, diferindo da classificação proposta inicialmente, em que seria um item de dificuldade intermediária. Seu parâmetro de discriminação foi muito baixo ( $a = 0,34$ ), gerando uma CCI pouco expressiva, que não diferencia muito as habilidades dos respondentes. Seu parâmetro de acerto casual foi zero ( $c = 0$ ), mas apesar disso, sua CII foi praticamente inexpressiva, mostrando que o item quase não tem informação, provavelmente devido à baixa discriminação. Dessa forma, esse item não é interessante pois não contribuirá na estimação de habilidades dos respondentes, sendo aconselhável a sua exclusão de um possível modelo de avaliação.

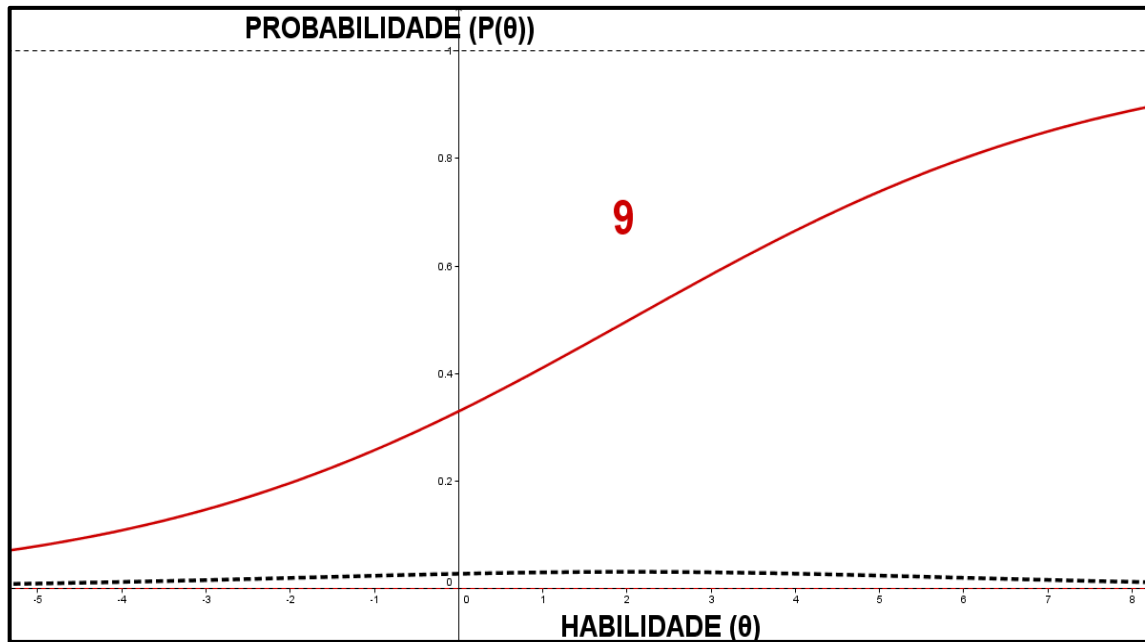




**Figura 22** - CCI e CII do item 8 do pré-teste.

### VIII. Item 9

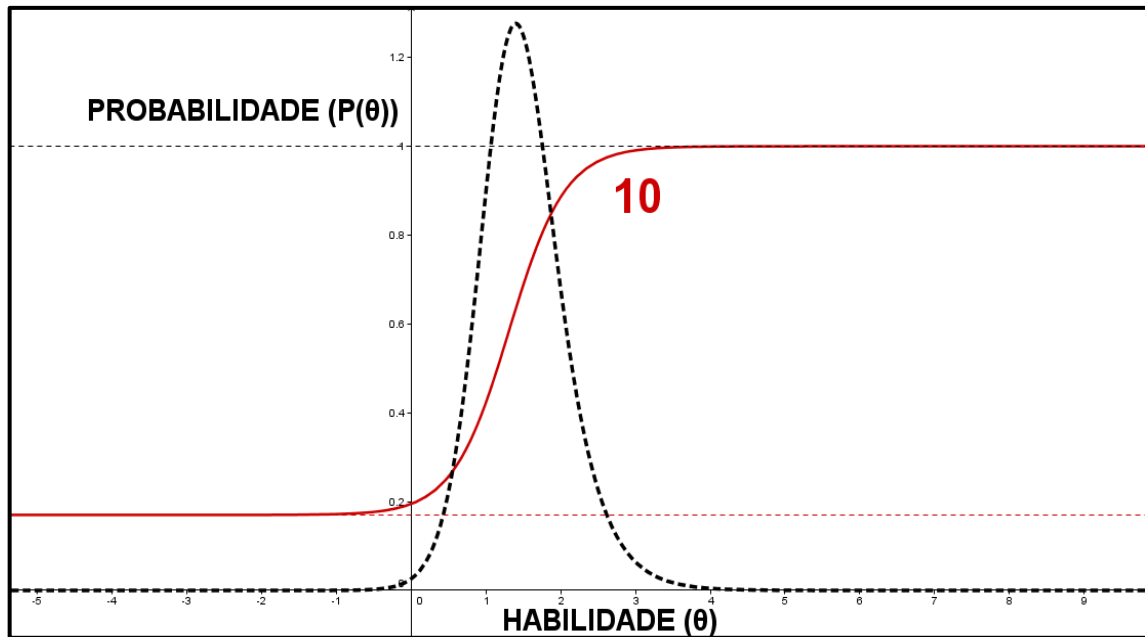
O item 9 teve um índice de acertos de 33,8%, que pode ser considerado um percentual intermediário quando comparado com os outros itens, dando a ideia que seria um item com dificuldade intermediária, concordando com a classificação proposta pelos pesquisadores inicialmente. Todavia, a estimação de seu parâmetro de dificuldade resultou num valor alto ( $b = 2,03$ ), fazendo com que seja considerado um item muito difícil. Seu parâmetro de discriminação apresenta um valor baixo ( $a = 0,35$ ) que se reflete na pouca inclinação de sua CCI e no comportamento quase inexpressivo da sua CII, que quase coincide com a assíntota inferior (Figura 23). Apesar de seu parâmetro de acerto casual ter sido zero ( $c = 0$ ), é um item que pouco discrimina habilidades e praticamente não traz informações, mostrando-se ruim para compor uma avaliação e por isso é sugerido que seja desconsiderado.



**Figura 23** - CCI e CII do item 9 do pré-teste.

#### **IX. Item 10:**

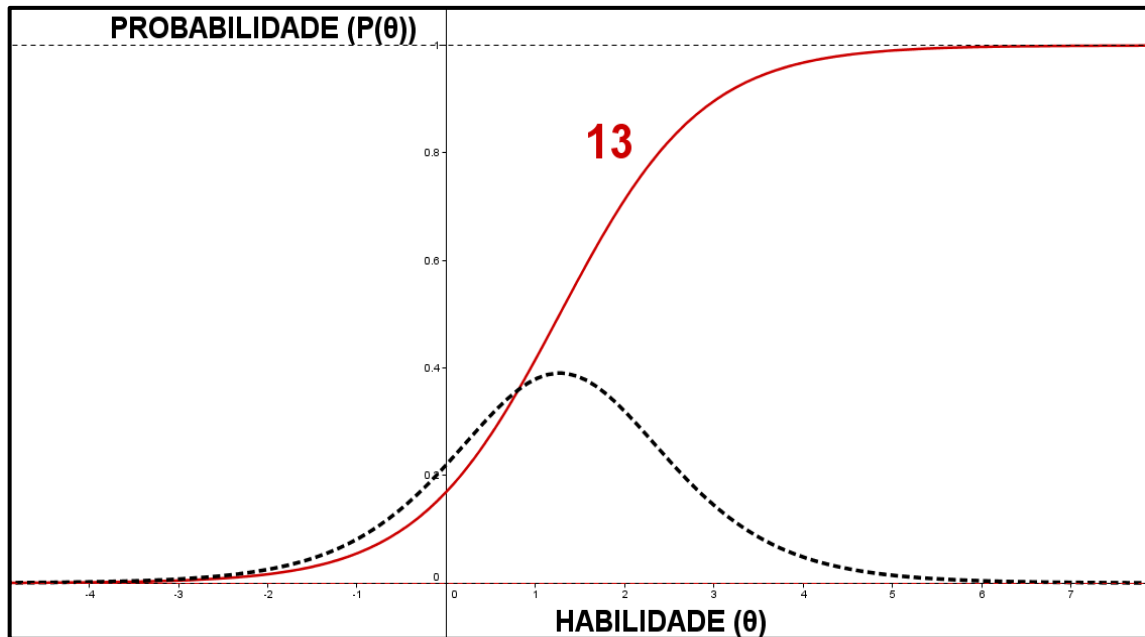
O item 10 apresentou um percentual de acerto de 28,7%, que pode ser considerado um pouco abaixo do valor intermediário, quando comparado com todos os outros, o que poderia aparentar uma maior dificuldade dos respondentes em resolver o item, considerado inicialmente com uma dificuldade acima da intermediária. A estimação do parâmetro de dificuldade gerou um valor que classifica o item como difícil ( $b = 1,31$ ), diferente da classificação inicial. Seu parâmetro de discriminação foi o mais alto de todo o teste ( $a = 2,66$ ), se mostrando o mais discriminativo de todos os itens e com uma curva bem mais inclinada que as dos outros. Apresentou um valor um pouco maior em relação aos demais itens com relação ao parâmetro de acerto casual ( $c = 0,17$ ), mas sua CII ainda foi muito expressiva, devido sua alta discriminação (Figura 24), mostrando que esse item traz muita informação. Dessa forma, esse item será utilizado para a constituição de avaliação.



**Figura 24** - CCI e CII do item 10 do pré-teste.

#### **X. Item 13:**

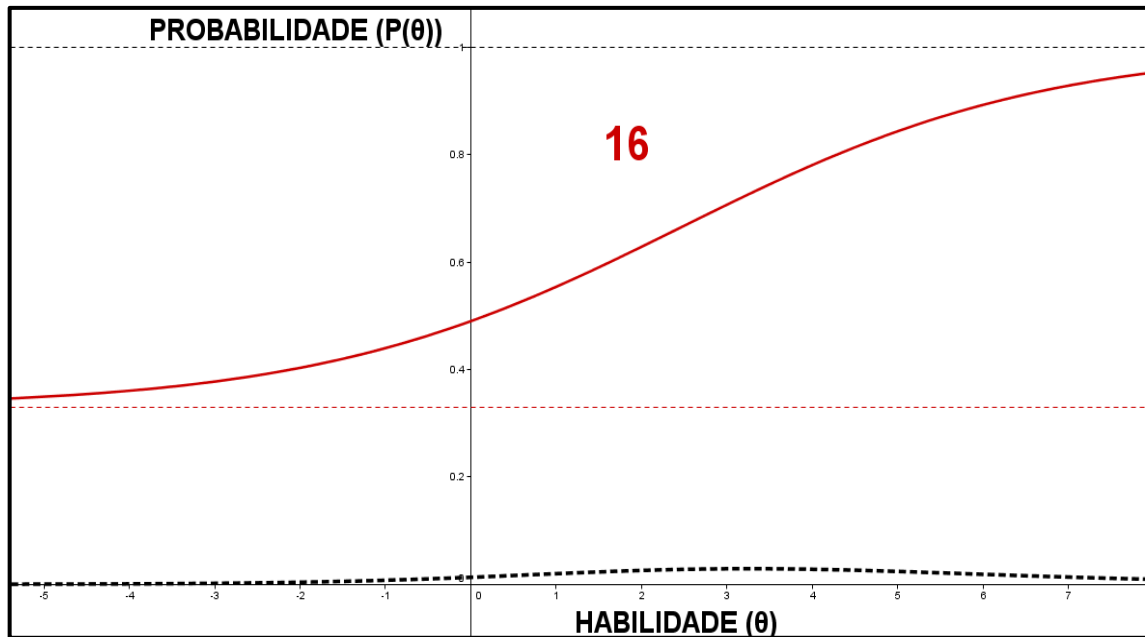
O item 13 obteve 22,69% de acertos, que numa análise inicial foi considerado baixo, quando comparado aos outros itens e poderia sugerir que o item era de alta dificuldade, ou pelo menos, acima da dificuldade intermediária. A estimação do parâmetro de dificuldade ( $b = 1,27$ ) levou o item a ser classificado como difícil, em desacordo com a classificação inicialmente proposta pelos pesquisadores. Seu parâmetro de discriminação apresentou um valor acima do intermediário ( $a = 1,25$ ), em relação aos outros, indicando que o item é bem discriminativo, como se pode ver em sua CCI (Figura 25), que se apresenta bastante inclinada. Seu parâmetro de acerto casual foi estimado em zero ( $c = 0$ ) e isso, junto com o parâmetro de discriminação, fizeram com que sua CII apresentasse um comportamento expressivo, indicando que o item traz bastante informação. A partir de todas essas características pode-se concluir que este é um bom item e poderia estar presente em uma avaliação.



**Figura 25** - CCI e CII do item 13 do pré-teste.

#### **XI. Item 16:**

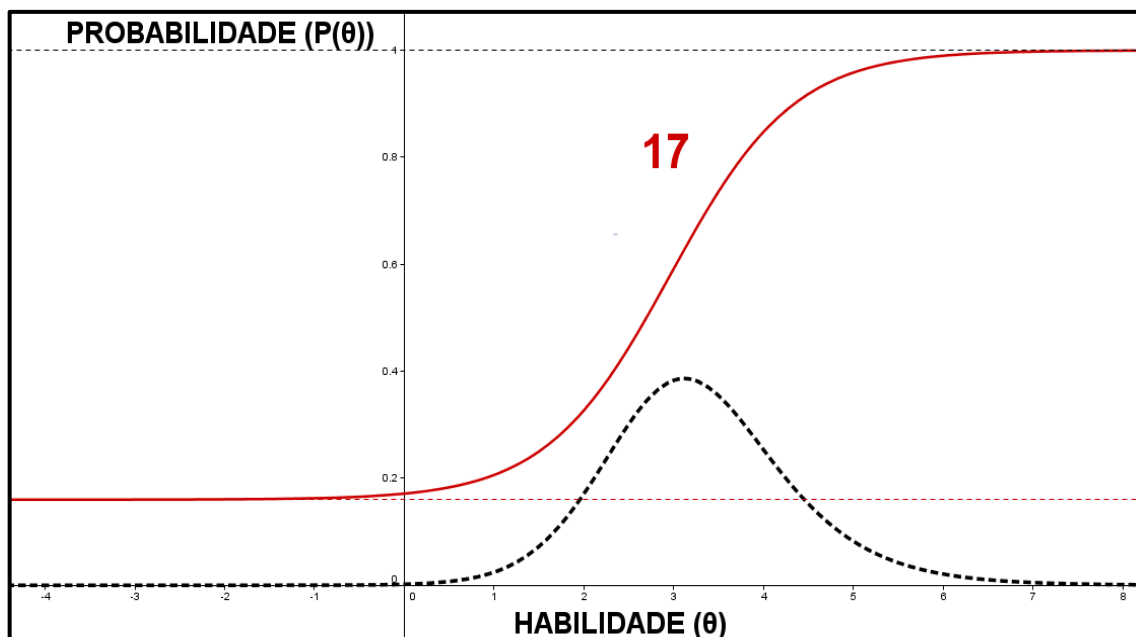
O item 16 obteve 49,54% de acertos entre os respondentes, que representa um dos maiores percentuais encontrados nesse teste (entre os itens que não foram excluídos inicialmente), dando a impressão de que esse item não apresentou muita dificuldade aos respondentes. Seu parâmetro de dificuldade foi um dos mais altos entre os itens que ainda estão sendo analisados ( $b = 2,46$ ), sendo classificado como uma questão muito difícil, diferente da classificação da Tabela 1, em que era colocado como item com dificuldade um pouco acima da dificuldade intermediária. Seu parâmetro de discriminação foi baixo ( $a = 0,47$ ), mostrando um item em que respondentes com diferentes graus de habilidades estão chegando à resposta correta (item pouco discriminativo), mostrando que o alto índice de acertos de um item pode não auxiliar na estimação de habilidades dos respondentes. Seu parâmetro de acerto casual foi o mais alto de todo o teste ( $c = 0,33$ ), que aliado aos outros parâmetros do item, deu origem às curvas da Figura 26: uma CCI mais deslocada para a direita (indicando maior dificuldade), com pouca inclinação (indicando sua baixa discriminação) e uma CII inexpressiva, indicando que o item traz pouquíssima informação (devido a baixa discriminação e alta probabilidade de acerto casual). Desse modo, esse item não é um bom item para a composição de uma avaliação, sendo desconsiderado.



**Figura 26 - CCI e CII do item 16 do pré-teste**

## **XII. Item 17:**

O item 17 apresentou 18,52% de acerto entre os respondentes, sendo o item com o menor percentual de acertos de todo o teste. A estimação de seu parâmetro dificuldade gerou o maior valor entre os itens que não foram eliminados inicialmente ( $b = 2,96$ ), que classifica esse item como muito difícil e faz com que suas curvas sejam mais deslocadas à direita no eixo das habilidades (Figura 27). Seu parâmetro de discriminação foi alto ( $a = 1,45$ ) e fez com que sua CCI ficasse mais inclinada, mostrando que é um item bem discriminativo. Apesar de seu parâmetro de acerto casual ter sido alto ( $c = 0,16$ ), que fez com que sua CII não atingisse valores maiores, ela ainda apresentou um comportamento expressivo (devido a alta discriminação) indicando um item com muita informação. Portanto esse item apresenta boas características e pode ser considerado na elaboração de uma avaliação.



**Figura 27** - CCI e CII do item 17 do pré-teste.

Com a finalização das análises de cada item, a partir da a estimação de seus parâmetros e construção de suas curvas, foi escolhido quais itens farão parte do modelo de avaliação que será construído. É importante notar que quanto maior a quantidade de itens no pré-teste, mais opções haverá para a elaboração da avaliação, procurando dessa forma utilizar somente os itens que tenham os melhores parâmetros, buscando assim, criar um instrumento capaz de gerar a melhor estimação das habilidades dos respondentes. Para escolha dos itens se dará preferência para as características já citadas anteriormente no texto, ou seja, alto grau de discriminação, baixa probabilidade de acerto casual, que fazem com que o item apresente alto grau de informação. O grau de dificuldade teve um caráter classificatório apenas, não sendo utilizado para descartar itens após as análises. Uma boa avaliação depende da escolha de itens bem calibrados, que poderão gerar resultados mais confiáveis em relação às habilidades que se pretende estimar.

Desse modo, para a elaboração da proposta de avaliação foram escolhidos sete itens do pré-teste que possuíam as características observadas no parágrafo anterior, sendo eles, os itens **1, 3, 4, 7, 10, 13 e 17**. Eles apresentam variado grau de dificuldade, que é uma característica importante no momento de se elaborar um teste, para que se possa obter uma melhor estimação das habilidades dos respondentes. Como neste caso está prevista a criação de um único modelo de avaliação, não haverá a necessidade da aplicação do processo de equalização, necessário somente quando há a intenção de se fazer modelos diferentes (mas com alguns itens comuns) para aplicação em vários grupos respondentes.

### 4.3 ESTIMAÇÃO DE HABILIDADES DOS RESPONDENTES A PARTIR DO R.

Outra análise possível de ser feita a partir dos resultados encontrados no pré-teste é a estimação das habilidades dos respondentes, através do software R. É importante salientar que essa estimação terá apenas um caráter ilustrativo, uma vez que esse processo só deveria ser feito através de um teste com itens calibrados (após o pré-teste e análise dos parâmetros encontrados), para que dessa forma fosse possível inclusive, gerar uma nota para os respondentes, de forma mais coerente e que represente as reais habilidades destes. Caso houvesse o interesse em estimar as habilidades de um grupo respondente de forma concreta, seria necessária a aplicação de novo teste, composto agora somente pelos itens que foram considerados adequados após o processo de calibração.

A partir dos resultados encontrados na resolução do pré-teste e utilizando o pacote *irt*, foi possível estimar as habilidades dos respondentes ( $\theta$ ) na escala logística, que está representada na Tabela 6.

**Tabela 6** - Estimativa das habilidades dos respondentes na escala logística (P = participante; H = habilidade)

P.	H.	P.	H.	P.	H.	P.	H.	P.	H.	P.	H.
1	0,7467	37	0,6530	73	-0,7692	109	-0,8376	145	-0,1165	181	-0,7774
2	1,2477	38	-0,0444	74	-0,8291	110	-0,2597	146	-0,7608	182	-0,4720
3	0,8418	39	-0,1434	75	-0,7746	111	0,6008	147	-0,5388	183	-0,0763
4	-0,1993	40	1,4740	76	0,0293	112	-0,5247	148	0,0201	184	-0,4700
5	0,6098	41	0,8007	77	-1,0155	113	-0,7715	149	-0,3603	185	-0,5837
6	1,2760	42	0,1618	78	-0,0444	114	-0,2079	150	-0,3772	186	-0,0253
7	1,8304	43	1,0059	79	0,3517	115	0,1554	151	-0,4755	187	0,2911
8	-0,3577	44	1,2421	80	0,1516	116	0,1402	152	-0,3180	188	-0,6688
9	0,1303	45	-0,0378	81	-0,5828	117	-0,7461	153	-0,6936	189	-0,8459
10	0,7748	46	1,2590	82	-1,0631	118	-1,1329	154	-0,4228	190	-0,2840
11	0,5751	47	0,7829	83	-0,9812	119	-0,7864	155	-0,0330	191	-0,5116
12	-0,2567	48	1,4505	84	-0,7503	120	-0,1661	156	-0,6846	192	-1,0279
13	0,5700	49	1,3013	85	-0,5568	121	-0,3213	157	-0,2398	193	0,0891
14	0,4470	50	0,9397	86	-0,0576	122	0,1159	158	-0,1358	194	-0,9101
15	0,3467	51	1,4259	87	0,3360	123	-0,0837	159	-0,8675	195	-0,9135
16	0,3380	52	1,1624	88	1,2184	124	0,1326	160	0,2992	196	-0,8670
17	0,1806	53	1,6215	89	0,4675	125	-0,4382	161	-0,2629	197	-1,0357
18	0,9658	54	0,5392	90	0,3525	126	-0,7315	162	-0,4019	198	-0,9006
19	-0,0906	55	0,2431	91	0,4524	127	-1,0270	163	0,3741	199	-0,6923
20	0,7791	56	0,5739	92	0,6285	128	-0,7722	164	-0,7844	200	0,0032
21	2,0440	57	0,4991	93	-0,3276	129	0,3829	165	0,5114	201	-1,0270
22	0,1380	58	0,6435	94	-0,7022	130	-0,8089	166	-0,7461	202	-0,0513
23	1,3497	59	-0,2523	95	0,3239	131	-0,4521	167	-0,2332	203	0,5249
24	0,3029	60	0,8844	96	-0,0494	132	-1,2114	168	-0,8101	204	-1,0775
25	-0,7172	61	1,2352	97	1,2676	133	-0,3229	169	-0,2238	205	-0,2235
26	1,2273	62	1,4328	98	-0,7183	134	-0,7845	170	-1,0537	206	-0,4302
27	-0,2049	63	1,8256	99	-0,5523	135	-0,0502	171	-0,5081	207	-0,4455
28	0,4332	64	0,9928	100	0,6584	136	-0,9938	172	-0,2079	208	-0,4150
29	0,5382	65	0,5586	101	-0,6011	137	-0,3091	173	-0,6321	209	0,5625

<b>30</b>	1,6299	<b>66</b>	0,6365	<b>102</b>	1,1569	<b>138</b>	-0,5826	<b>174</b>	-0,2853	<b>210</b>	-0,9099
<b>31</b>	1,0244	<b>67</b>	0,0572	<b>103</b>	-0,7183	<b>139</b>	-0,4237	<b>175</b>	0,1551	<b>211</b>	-0,7857
<b>32</b>	-0,1930	<b>68</b>	0,8545	<b>104</b>	-0,1957	<b>140</b>	-0,9806	<b>176</b>	0,4783	<b>212</b>	-0,8033
<b>33</b>	-0,7189	<b>69</b>	0,6221	<b>105</b>	-0,7420	<b>141</b>	-0,0711	<b>177</b>	0,9256	<b>213</b>	-1,0364
<b>34</b>	-0,0088	<b>70</b>	1,0232	<b>106</b>	0,8512	<b>142</b>	-1,0780	<b>178</b>	0,4778	<b>214</b>	-0,3067
<b>35</b>	0,3077	<b>71</b>	0,6718	<b>107</b>	-1,0680	<b>143</b>	0,0004	<b>179</b>	-0,8459	<b>215</b>	-0,4202
<b>36</b>	1,0874	<b>72</b>	2,1893	<b>108</b>	0,4060	<b>144</b>	-0,8637	<b>180</b>	-0,7516	<b>216</b>	-0,6017

As habilidades estimadas através da TRI podem assumir quaisquer valores reais, diferente da TCT, em que as habilidades eram apenas uma soma de pontos obtidos através dos itens respondidos corretamente. Ao observar as informações contidas na Tabela 6, que representam as habilidades dos respondentes estimadas na escala logística, encontram-se números positivos e negativos, que podem causar dúvidas quanto a seus significados para pesquisadores com menos experiência em Estatística, principalmente quando se tem em mente um respondente com habilidade negativa. Por essa razão é necessário analisar com calma o significado desses resultados e, caso haja a necessidade de se gerar uma nota para os respondentes, é recomendável adotar uma nova escala que consiga representar de forma mais clara esse resultado.

Ao fazer a estimação das habilidades é necessário definir uma escala padrão, que dará significado para os números que foram encontrados. Essa escala deve ser de conhecimento de quem fará a análise dos resultados obtidos, ou não será possível concluir como foi o desempenho dos respondentes. Para essa definição utiliza-se um valor para ser o ponto central, ou origem da escala (que será o valor médio das habilidades dos respondentes) e outro que será sua unidade de medida (o desvio-padrão). Geralmente o modelo adotado (inclusive no caso da Tabela 6) é o padrão  $(0,1)$ , em que se considera o valor médio igual a  $0$  e o desvio padrão igual a  $1$ . Logo, os valores da tabela não estão representando notas, mas o quanto a habilidade do respondente se aproxima (ou se afasta) da habilidade média, em quantidade de desvios-padrões. Para melhor explicação, considere um respondente qualquer, por exemplo, o de número 193, que teve sua habilidade estimada em  $0,0891$ . Este valor significa que a habilidade do respondente está  $0,0891$  desvios-padrões acima da média das habilidades.

Dessa forma se estabelece uma relação de ordem entre os valores encontrados para as habilidades, baseado no valor definido como habilidade média. Assim, uma habilidade negativa representa que aquele respondente está abaixo da habilidade média, sem estar diretamente ligada à nota. Para a definição de uma nota é necessário fazer uma mudança de escala, alterando o valor médio e o desvio-padrão, definindo valores que sejam mais condizentes com a realidade que se deseja avaliar. Todas as avaliações que atualmente trabalham com a TRI fazem essa mudança na escala, com o intuito de apresentar resultados



que sejam mais compreensíveis aos respondentes e às pessoas que pretendem utilizar as informações encontradas a partir do teste.

O exemplo mais comum de avaliação no modelo da TRI existente hoje no Brasil é o Exame Nacional do Ensino Médio (Enem), ofertado anualmente pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira). Inicialmente era uma avaliação diagnóstica, com o intuito de avaliar a qualidade do Ensino Médio e das escolas que oferecem essa modalidade. Ao longo dos anos, outras funções foram agregadas a essa avaliação, sendo as principais, servir de instrumento de certificação do Ensino Médio para pessoas que estão fora da idade indicada para essa modalidade (pessoas acima de 18 anos); classificar e selecionar pessoas para ingressar em instituições de ensino técnico ou ensino superior, públicas (Sisutec<sup>4</sup> ou Sisu<sup>5</sup>) ou privadas e, classificar pessoas em programas governamentais, como os que oferecem bolsas de estudos em universidades privadas (PROUNI<sup>6</sup>).

Para todas as funções listadas anteriormente é necessária a atribuição de notas aos respondentes, para que seja possível efetuar sua classificação, ou classificação das escolas (a partir das notas dos alunos). A classificação na escala logística é menos perceptiva e dessa forma, o INEP adota uma métrica diferenciada para o Enem, considerando 500 como valor médio de habilidade (ponto central, ou origem) e 100 como desvio-padrão (unidade de medida), gerando as notas que são apresentadas aos respondentes e todas as outras pessoas interessadas. É importante salientar que ao gerar essas notas houve apenas uma mudança de escala, sem alterar as relações de ordem definidas a partir das análises iniciais e dos resultados obtidos na escala logística.

Com o intuito de mostrar como seriam as notas dos respondentes dessa pesquisa dentro da escala padrão do Enem, foi utilizado o pacote *CTT* do software R para que os dados da Tabela 6 fossem convertidos a essa métrica (média 500 e desvio-padrão 100) e eles estão apresentados na Tabela 7.

---

<sup>4</sup> **O Sistema de Seleção Unificada da Educação Profissional e Tecnológica** (Sisutec) é utilizado para selecionar alunos que queiram cursar educação profissional e tecnológica em instituições públicas ou privadas, que aderiram a esse sistema. Ele utiliza as notas do Enem como critério de classificação dos candidatos.

<sup>5</sup> **O Sistema de Seleção Unificada** (Sisu) é utilizado para selecionar alunos para o nível superior em instituições públicas, a partir do desempenho no Enem.

<sup>6</sup> **Prouni** é um programa criado pelo governo federal que concede bolsas de estudos parciais (50%) ou integrais em instituições privadas de ensino superior para candidatos que tenham feito Enem e obtido pontuação mínima de 450, além de não ter tirado zero na redação.

**Tabela 7** - Estimativa das habilidades dos respondentes na escala adotada pelo Enem (500,100)

<b>P.</b>	<b>H.</b>	<b>P.</b>	<b>H.</b>	<b>P.</b>	<b>H.</b>	<b>P.</b>	<b>H.</b>	<b>P.</b>	<b>H.</b>	<b>P.</b>	<b>H.</b>
<b>1</b>	599,33	<b>37</b>	586,86	<b>73</b>	397,72	<b>109</b>	399,62	<b>145</b>	385,15	<b>181</b>	396,63
<b>2</b>	665,96	<b>38</b>	494,11	<b>74</b>	389,76	<b>110</b>	465,48	<b>146</b>	484,53	<b>182</b>	437,25
<b>3</b>	611,98	<b>39</b>	480,95	<b>75</b>	397,01	<b>111</b>	579,92	<b>147</b>	398,84	<b>183</b>	489,88
<b>4</b>	473,51	<b>40</b>	696,05	<b>76</b>	503,91	<b>112</b>	430,24	<b>148</b>	428,36	<b>184</b>	437,51
<b>5</b>	581,11	<b>41</b>	606,51	<b>77</b>	364,96	<b>113</b>	397,41	<b>149</b>	502,70	<b>185</b>	422,40
<b>6</b>	669,71	<b>42</b>	521,54	<b>78</b>	494,12	<b>114</b>	472,37	<b>150</b>	452,11	<b>186</b>	496,66
<b>7</b>	743,45	<b>43</b>	633,80	<b>79</b>	546,79	<b>115</b>	520,69	<b>151</b>	449,86	<b>187</b>	538,74
<b>8</b>	452,45	<b>44</b>	665,22	<b>80</b>	520,18	<b>116</b>	518,67	<b>152</b>	436,79	<b>188</b>	411,07
<b>9</b>	517,35	<b>45</b>	494,99	<b>81</b>	422,51	<b>117</b>	400,80	<b>153</b>	457,73	<b>189</b>	387,52
<b>10</b>	603,07	<b>46</b>	667,46	<b>82</b>	358,64	<b>118</b>	349,36	<b>154</b>	407,77	<b>190</b>	462,24
<b>11</b>	576,50	<b>47</b>	604,14	<b>83</b>	369,53	<b>119</b>	395,44	<b>155</b>	443,79	<b>191</b>	431,99
<b>12</b>	465,88	<b>48</b>	692,93	<b>84</b>	400,24	<b>120</b>	477,93	<b>156</b>	495,63	<b>192</b>	363,31
<b>13</b>	575,82	<b>49</b>	673,08	<b>85</b>	425,97	<b>121</b>	457,29	<b>157</b>	408,98	<b>193</b>	511,87
<b>14</b>	559,47	<b>50</b>	625,00	<b>86</b>	492,35	<b>122</b>	515,43	<b>158</b>	468,13	<b>194</b>	378,99
<b>15</b>	546,13	<b>51</b>	689,65	<b>87</b>	544,71	<b>123</b>	488,89	<b>159</b>	481,96	<b>195</b>	378,53
<b>16</b>	544,97	<b>52</b>	654,61	<b>88</b>	662,06	<b>124</b>	517,65	<b>160</b>	384,64	<b>196</b>	384,71
<b>17</b>	524,04	<b>53</b>	715,67	<b>89</b>	562,19	<b>125</b>	441,75	<b>161</b>	539,81	<b>197</b>	362,27
<b>18</b>	628,47	<b>54</b>	571,73	<b>90</b>	546,90	<b>126</b>	402,74	<b>162</b>	465,06	<b>198</b>	380,25
<b>19</b>	487,97	<b>55</b>	532,35	<b>91</b>	560,19	<b>127</b>	363,44	<b>163</b>	446,57	<b>199</b>	407,95
<b>20</b>	603,63	<b>56</b>	576,34	<b>92</b>	583,60	<b>128</b>	397,33	<b>164</b>	395,70	<b>200</b>	500,44
<b>21</b>	771,86	<b>57</b>	566,40	<b>93</b>	456,46	<b>129</b>	550,94	<b>165</b>	568,04	<b>201</b>	363,44
<b>22</b>	518,37	<b>58</b>	585,60	<b>94</b>	406,63	<b>130</b>	392,45	<b>166</b>	400,80	<b>202</b>	493,20
<b>23</b>	679,52	<b>59</b>	466,46	<b>95</b>	543,10	<b>131</b>	439,89	<b>167</b>	469,00	<b>203</b>	569,83
<b>24</b>	540,30	<b>60</b>	617,63	<b>96</b>	493,45	<b>132</b>	338,91	<b>168</b>	392,28	<b>204</b>	356,72
<b>25</b>	404,63	<b>61</b>	664,29	<b>97</b>	668,60	<b>133</b>	457,08	<b>169</b>	470,26	<b>205</b>	470,30
<b>26</b>	663,25	<b>62</b>	690,57	<b>98</b>	404,49	<b>134</b>	395,68	<b>170</b>	359,89	<b>206</b>	442,81
<b>27</b>	472,76	<b>63</b>	742,82	<b>99</b>	426,56	<b>135</b>	493,34	<b>171</b>	432,44	<b>207</b>	440,77
<b>28</b>	557,64	<b>64</b>	632,05	<b>100</b>	587,59	<b>136</b>	367,86	<b>172</b>	472,37	<b>208</b>	444,83
<b>29</b>	571,60	<b>65</b>	574,31	<b>101</b>	420,08	<b>137</b>	458,92	<b>173</b>	415,95	<b>209</b>	574,83
<b>30</b>	716,79	<b>66</b>	584,68	<b>102</b>	653,88	<b>138</b>	422,53	<b>174</b>	462,08	<b>210</b>	379,02
<b>31</b>	636,25	<b>67</b>	507,63	<b>103</b>	404,49	<b>139</b>	443,67	<b>175</b>	520,64	<b>211</b>	395,52
<b>32</b>	474,36	<b>68</b>	613,66	<b>104</b>	474,00	<b>140</b>	369,61	<b>176</b>	563,63	<b>212</b>	393,19
<b>33</b>	404,41	<b>69</b>	582,76	<b>105</b>	401,34	<b>141</b>	490,56	<b>177</b>	623,12	<b>213</b>	362,18
<b>34</b>	498,86	<b>70</b>	636,10	<b>106</b>	613,22	<b>142</b>	356,65	<b>178</b>	563,57	<b>214</b>	459,23
<b>35</b>	540,94	<b>71</b>	589,37	<b>107</b>	357,98	<b>143</b>	500,07	<b>179</b>	387,52	<b>215</b>	444,14
<b>36</b>	644,64	<b>72</b>	791,18	<b>108</b>	554,02	<b>144</b>	385,15	<b>180</b>	400,07	<b>216</b>	420,00

Os resultados apresentados na Tabela 7 representam as notas dos respondentes do pré-teste dentro da métrica (500, 100), mas elas poderiam ser descritas em qualquer outra métrica desejada, conforme o tipo de avaliação que está sendo feito ou os critérios utilizados. É importante reforçar que qualquer tipo de análise ou, qualquer situação em que essas notas forem utilizadas é necessário que se tenha o conhecimento de qual a métrica utilizada e, conseqüentemente, qual o significado dos valores encontrados. Um candidato que tenha apresentado nota superior a 500 representa uma habilidade acima do valor médio estipulado, assim como um que tenha nota abaixo desse valor representa uma habilidade abaixo do valor médio.

Um equívoco muito comum envolvendo testes no modelo da TRI é pensar que as notas obtidas estão diretamente relacionadas ao total de acertos e erros do teste (os chamados escores brutos), geralmente usados em testes feitos no modelo clássico. Muitas vezes ao verificar as notas, os respondentes querem relacioná-las à quantidade de itens que acertaram e procuram inclusive, estimar valores para cada item que acertaram, chegando a acreditar que foram atribuídos valores diferentes aos itens ou valores não inteiros. Nessa mesma linha de pensamento, muitos respondentes ficam intrigados ao perceberem que pessoas com a mesma quantidade de acertos tiveram notas diferentes ao verificar o resultado oficial da avaliação.

Essas confusões são geradas no momento em que se tenta analisar um teste feito no modelo TRI como se fosse no modelo clássico. Na TRI a nota não é gerada a partir do escore bruto e sim, a partir da análise de cada item e da solução dada pelo respondente, sendo muito plausível existir sujeitos com quantidades de acertos brutos iguais e notas finais diferentes, pois o modelo da TRI é sensível a quais questões foram respondidas corretamente. Se dois respondentes tiverem o mesmo escore bruto, mas um deles acertou somente questões consideradas fáceis e médias, enquanto o outro acertou questões médias e difíceis, o segundo terá uma nota maior que o primeiro. Outra influência que não se pode ignorar é a probabilidade do acerto casual, que também altera as notas finais, fazendo com que fiquem menores caso essa probabilidade seja muito alta.

Assim, é muito importante observar que os valores apresentados na Tabela 7 não são somente notas, somatórios de pontos obtidos através de respostas corretas, mas são valores que representam, dentro de uma escala determinada, a habilidade de um sujeito em relação ao conteúdo abordado dentro do teste que lhe foi aplicado, levando-se em consideração as características de cada item (vale lembrar que os valores dessa tabela são meramente ilustrativos, sendo importante a discussão sobre os métodos utilizados para se chegar nesses valores).

#### **4.4 A TRI DENTRO DA REALIDADE ESCOLAR: UMA POSSIBILIDADE.**

Nos últimos anos tem-se visto uma popularização do uso da TRI, principalmente após sua adoção como modelo de avaliação por vários órgãos governamentais, ligados à área da educação, com o intuito principal de verificar os conhecimentos gerais dos alunos e, a partir disso, classificá-lo ou classificar sua instituição de ensino, para as mais variadas finalidades.

A adoção desse novo modelo de avaliação abre uma série de possibilidades que não

existiam no modelo tradicional, focado somente na quantidade de itens respondidos corretamente, que eram convertidos num número (a famosa nota) que classificava (ou pode-se dizer, marcava) o respondente dentro de uma escala definida previamente. O foco estava no número, sem se preocupar com a origem dele, ou como ele foi formado, pensando apenas que quanto maior melhor, sem levar em consideração se esse número foi gerado a partir das reais habilidades do respondente ou simplesmente pelo acaso de ter respondido corretamente.

O teste clássico ainda é muito adotado por professores dentro das escolas, ou mesmo em outras áreas, como os famosos processos seletivos ou concursos, devido a sua praticidade para a obtenção de notas, sem a necessidade de se utilizar complexos modelos matemáticos ou softwares avançados para análises. Entretanto, os resultados obtidos a partir desse modelo são cada vez mais questionados quando se pretende relacioná-los com o grau de habilidade do respondente, mostrando que o uso da TRI tende a aumentar cada vez mais.

Em teoria, um instrumento de avaliação feito seguindo o modelo da TRI seria mais justo que no modelo tradicional, uma vez que não se analisa somente o total de acertos do respondente, mas se leva em consideração quais foram esses acertos, considerando o grau de dificuldade dos itens marcados corretamente e dos que foram marcados errados e considerando também, a possibilidade da ocorrência do acerto casual. Esse instrumento de avaliação é montado a partir de itens previamente testados, seguindo um padrão de organização em que os itens apresentam diversos graus de dificuldade, com o intuito de se encontrar diversos graus de habilidades.

No modelo clássico utiliza-se os acertos para gerar uma nota e a partir desta, os respondentes são classificados em relação a sua habilidade, enquanto na TRI utiliza-se todos os itens (certos e errados) para estimar suas habilidades, para então gerar a nota que o classificará. Assim, a utilização desse novo modelo desperta uma série de dúvidas e faz com que se questionem os caminhos adotados até então nos processos de avaliação, levantando várias questões: Qual o objetivo da avaliação, afinal? A avaliação serve para verificar os conhecimentos (ou habilidades) dos alunos ou é apenas um instrumento classificador, que visa somente a atribuição de uma nota (transformando o respondente em um número)?

Ao se discutir sobre o uso de um determinado modelo de avaliação, acaba-se retomando antigas discussões sobre qual o objetivo, em teoria, do processo de avaliação e, o que realmente ocorre na prática, dentro das salas de aulas ou nos processos seletivos existentes. Como se sabe, a avaliação tem o objetivo de verificar os conhecimentos desenvolvidos por um sujeito em uma determinada situação, a partir da aplicação de algum instrumento que gera um resultado, que será convertido numa classificação que indica a

habilidade do sujeito. Todavia, a prática se afasta um pouco da teoria e o que vemos são avaliações que servem somente para atribuir notas aos sujeitos, sem uma análise concreta de suas habilidades, considerando que aqueles que apresentam a mesma nota possuem a mesma habilidade e aqueles que não apresentam resultado satisfatório em uma avaliação não possuem habilidade. As análises são bem simplistas, sempre focadas em quantidades e nunca em qualidade, a partir de instrumentos que muitas vezes nem foram bem elaborados e não são capazes de fornecer resultados confiáveis.

O uso da TRI para as avaliações em larga escala tem crescido cada vez mais e apresentado resultados favoráveis, mostrando que nesse modelo, o objetivo principal do processo avaliativo tem se cumprido de forma satisfatória. Para a aplicação desse modelo dentro da escola tradicional seria necessário algumas adaptações e uma preparação, procurando facilitar sua utilização pelos professores. Essa preparação começa pelo método de elaboração dos itens que seriam aplicados, para que as avaliações sejam formadas por itens que consigam ser analisados dentro do modelo da TRI, procurando variar o grau de dificuldade e aumentar o grau de discriminação. Esse passo inicial é primordial para a adoção da TRI como um modelo de avaliação escolar, uma vez que não seria possível utilizar o modelo para itens que não tivessem as características básicas necessárias, pois eles seriam, provavelmente, eliminados durante o processo de calibração, o que iria gerar a necessidade de se elaborar mais itens para repor aqueles que foram descartados.

Outra condição fundamental para a adoção do modelo seria a capacitação em relação a utilização dos softwares necessários para fazer a calibração dos itens (com a estimação de parâmetros) e a estimação das habilidades (e sua conversão para a escala de notas adotada). Atualmente, muitas escolas possuem laboratórios de informática e esse processo poderia ocorrer no próprio ambiente escolar, usando melhor os espaços existentes.

Uma possibilidade muito interessante para a adoção do método dentro das escolas seria a criação de equipes responsáveis pela elaboração, correção e análises das avaliações, dentro do modelo da TRI. Uma equipe seria treinada no processo de elaboração dos itens, seguindo o padrão necessário ao modelo e agiria como multiplicador junto aos outros professores, para que todos possam aprender como se dá a elaboração do item e assim, contribuir fornecendo-os. Seria necessária a realização do teste para verificar se esses itens estão dentro de um padrão ideal de avaliação da TRI, sendo mais interessante fazê-lo fora da escola, mas com indivíduos pertencentes a grupos semelhantes aos que se deseja avaliar, para assim conseguir uma boa estimação dos parâmetros. Outra equipe seria responsável pela correção dessas avaliações, a partir de um gabarito previamente elaborado e então,

transformar esses resultados em tabelas binárias, que serão usadas pelo software para efetuar as análises. A última equipe seria responsável pela análise dos resultados através do software e então fornecer as notas esperadas, dentro da métrica adotada pela escola.

O processo de elaboração de itens é difícil no início, devido às pessoas não estarem acostumadas a seguirem o padrão necessário, mas num segundo momento, torna-se mais tranquilo, quando os elaboradores acostumam-se com a metodologia necessária. Essa ação em longo prazo irá gerar o chamado banco de itens – banco de dados formados por itens no modelo da TRI, para serem usados na criação de uma avaliação – que poderia ser organizado por disciplinas, séries, grau de dificuldade ou qualquer outra característica de interesse, criando-se uma fonte contínua de fornecimento de itens para a elaboração de avaliações, facilitando muito esse processo, principalmente porque esses itens já estarão calibrados e prontos para serem inseridos dentro de avaliações.

Atualmente, as chamadas avaliações externas em larga escala (Enem, Saesp<sup>7</sup>, Prova Brasil<sup>8</sup>) adotam o modelo da TRI e se propõe a servirem como diagnóstico dos processos educacionais. Entretanto, são avaliações que muitas vezes não abrangem todos os alunos, devido sua periodicidade e seu caráter amostral (a Prova Brasil, por exemplo, só é aplicada aos alunos do 5º e 9º ano do Ensino Fundamental, a cada dois anos), além de ocorrerem apenas uma vez durante o ano, gerando uma grande expectativa (e até nervosismo) nos alunos que farão a avaliação e, nas próprias escolas, uma vez que esse resultado tem um caráter classificatório, definindo quais serão as chamadas “melhores escolas” ou “escolas ruins”. Toda essa pressão pode afetar negativamente os avaliados, alterando resultados e não mostrando, muitas vezes, as reais habilidades que eles possuem. Se, no entanto, esse modelo de avaliação estivesse dentro do ambiente escolar, os alunos já estariam acostumados a trabalhar com ele e não se sentiriam tão pressionados a obter bons resultados, ou não seriam surpreendidos por um modelo de prova diferente do que estão acostumados a trabalhar.

A adoção desse método dentro da escola, para avaliação das disciplinas, seria um processo longo que precisaria ser feito com calma, necessitando de um grande número de

---

<sup>7</sup> **Sistema de Avaliação do Rendimento Escolar de São Paulo** (SARESP) é uma avaliação externa da Educação Básica, realizada desde 1996 pela Secretaria da Educação do Estado de São Paulo – SEE/SP, com o objetivo de obter informações sobre a situação da escolaridade básica na rede pública de ensino paulista, visando orientar os gestores do ensino no monitoramento das políticas voltadas para a melhoria da qualidade educacional.

<sup>8</sup> **A Avaliação Nacional do Rendimento Escolar** (Anresc), conhecida como Prova Brasil, é uma avaliação externa aplicada aos alunos do 5º e 9º ano do Ensino Fundamental das escolas públicas brasileiras, com o objetivo de auxiliar os governantes nas decisões e no direcionamento de recursos técnicos e financeiros, assim como a comunidade escolar, no estabelecimento de metas e na implantação de ações pedagógicas e administrativas, visando à melhoria da qualidade do ensino.

pessoas envolvidas para que dê certo, sendo esse o maior empecilho para essa ação. Outro fator de dificuldade seria a realização do teste para a calibração dos itens, de forma que todos esses itens tenham seus parâmetros estimados antes de tomarem parte dentro de um modelo de avaliação. Para isso, há a necessidade de se ter um número grande de itens (sempre existe a possibilidade de eliminar itens por não apresentarem bons parâmetros), para dessa forma, sempre haver itens suficientes para a elaboração de testes.

Mesmo com algumas dificuldades para sua implementação em um ambiente escolar, as respostas obtidas com o uso dos modelos da TRI podem ser muito importantes dentro do processo educacional, por trazerem resultados mais seguros acerca das habilidades desenvolvidas pelos alunos. Isso é extremamente importante durante o processo de ensino-aprendizagem, para que se possa verificar se as estratégias e métodos adotados pelo professor estão gerando os resultados esperados ou, se é necessário uma mudança nesse processo, para que se possa chegar cada vez mais próximo dos objetivos do ensino.

## 5 CONSIDERAÇÕES FINAIS

Ao finalizar a pesquisa, observa-se que a TRI apresenta muitas aplicações em diferentes áreas do conhecimento, sendo a educação um campo em que ela já vem sendo adotada de forma cada vez mais intensa, através das avaliações em larga escala, utilizadas para estimar as habilidades dos alunos e, através dessas, classificá-los ou classificar as suas instituições de ensino, com as mais diferentes finalidades. A teoria clássica ainda é muito adotada no ambiente escolar, muitas vezes por ser mais simples de se trabalhar ou, até mesmo por desconhecimento de outro modelo para elaboração das avaliações e por isso, é de extrema importância que ocorram discussões e debates sobre a TRI, procurando dessa forma, levar para dentro do ambiente escolar um modelo que já está próximo a ele.

Na perspectiva de aproximar o modelo da realidade escolar, a aplicação do teste envolvendo o tema “Tratamento da Informação” foi muito importante por mostrar quais os passos necessários e, quais não deveriam ser tomados, para a utilização da teoria dentro da escola. Em relação ao pré-teste, ficou clara a necessidade de se utilizar mais itens, ou se fazer mais testes, pois após a calibração sempre ocorre a exclusão de itens, diminuindo as opções de escolha para compor um modelo de avaliação. Além disso, com a existência de vários itens calibrados é possível a criação de um banco de itens, armazenando material diversificado (por exemplo, com vários graus de dificuldade) que poderá ser utilizado em outras avaliações, facilitando o processo de elaboração destas. Por conta do teste aplicado possuir poucos itens (apenas 17 itens válidos), não ocorreu grande diversidade no grau de dificuldade, que pode prejudicar a estimação das habilidades dos respondentes,

A estimação de habilidades feita teve um caráter meramente ilustrativo, uma vez que partiu dos resultados obtidos na aplicação do pré-teste, e não de itens calibrados previamente. Todavia, a informação apresentada nas análises foi muito importante para mostrar qual o comportamento esperado dos resultados e qual deve ser a interpretação destes, principalmente dissociando a ideia de nota obtida pelo total de pontos (score bruto) de nota obtida na análise via TRI. Esta leva em consideração vários fatores desprezados pela teoria clássica, como o grau de dificuldade das questões que foram marcadas corretamente ou a possibilidade do acerto casual do respondente, fazendo toda a diferença no cálculo da chamada “nota final”, que representará de forma mais concreta, a habilidade do respondente.

Um modelo de avaliação que não se preocupa somente com o total de acertos que um respondente apresenta tende a trazer muito mais informação sobre as reais habilidades que ele



possui, de forma a contribuir de forma mais intensa para as análises dos processos de aprendizagem. O respondente pode ser submetido a uma avaliação diagnóstica no início do processo de aprendizagem e ao final dele passar por outra e, esses resultados podem ser comparados de forma mais clara e segura, representando qual foi a evolução de suas habilidades ou, quais dificuldades ele tem apresentado. Além disso, o modelo permite ainda a comparação entre indivíduos diferentes e isso pode ser usado para avaliar um grupo respondente dentro de uma determinada instituição, mostrando como o processo de ensino-aprendizagem está se desenvolvendo entre os alunos.

Assim, o modelo da TRI está condizente com as atuais discussões envolvendo avaliação, dentro da área educacional, restando apenas a sua aplicação mais próxima ao ambiente escolar. Isso significa não estar preocupado somente com classificações em vestibulares, em rankings internacionais, federais ou estaduais, mas estar focado no objetivo principal de uma avaliação, que é observar se o aluno conseguiu adquirir, ou desenvolver, um determinado conceito trabalhado na escola e, caso isso não tenha ocorrido, quais as estratégias necessárias para resolver esse problema, de forma que seja possível a formação de um aluno mais capacitado que tenha possibilidade de se desenvolver dentro de uma sociedade que a cada dia dá menos espaço para as pessoas que não conseguem se integrar em seu modelo já definido.

## REFERÊNCIAS

- ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria da Resposta ao Item: conceitos e aplicações.** In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 14., 2000., Caxambú, MG: Associação Brasileira de Estatística, 2000.
- ANDRIOLA, W. B. **Psicometria Moderna: características e tendências. Estudo em Avaliação Educacional,** São Paulo, v. 20, n. 43, agosto 2009.
- ANJOS, A.; ANDRADE, D. F. **Teoria de Resposta ao Item com o uso do R.** In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 20., 2012., João Pessoa, PB: Associação Brasileira de Estatística, 2012.
- ARAÚJO, E. A. C.; ANDRADE, D. F.; BORTOLOTTI, S. L. V. **Teoria da Resposta ao Item. Revista da Escola de Enfermagem da USP,** São Paulo, v. 43, n. spe, Dec. 2009.
- BRAGION, M. L. L. **Um modelo de Teoria de Teoria de Resposta ao Item para os dados do vestibular 2006-2 da UFLA.** 2007. 88 p. Dissertação (Mestrado) – Universidade Federal de Lavras, Minas Gerais. 2007
- Comissão Permanente para Vestibulares. **Exercícios do pré-teste.** Disponível em: <<https://www.comvest.unicamp.br>> Acesso em: 01 fev. 2014
- GeoGebra. **Sobre.** Disponível em: <[www.geogebra.org](http://www.geogebra.org)> Acesso em: 01 fev. 2014.
- PASQUALI, L., PRIMI, R. **Fundamentos da Teoria da Resposta ao Item –TRI. Avaliação Psicológica,** Porto Alegre, v. 2, n. 2, 2003.
- PEREIRA, V. R. **Métodos alternativos no critério Brasil para construção de indicadores sócio-econômico: teoria da resposta ao item.** 2004. 103 p. Dissertação (Mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro. 2004
- Professor Cardy. **Exercícios do pré-teste.** Disponível em: <<http://www.profcardy.com/exercicios>> Acesso em: 01 fev. 2014
- Prouni -. **Página Inicial.** Disponível em: <<http://siteprouni.mec.gov.br/>> Acesso em: 01 fev. 2014
- R Development Core Team. **R: A language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing. Disponível em <<http://www.r-project.org>>. Acesso em 01 fev. 2014

Saeb: **Histórico do Saeb**. Disponível em: <<http://portal.inep.gov.br/web/saeb/historico>>

Acesso em: 01 fev. 2014

Saresp 2013 - **Apresentação**. Disponível em: <<http://saresp.fde.sp.gov.br/2013>> Acesso em: 01 fev. 2014

SARTES, L. M. A.; SOUZA-FORMIGONI, M. L. O. Avanços na psicometria: da Teoria Clássica dos Testes à Teoria de Resposta ao Item. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 26, n. 2, 2013.

Sisu - **Página inicial**. Disponível em: <<http://sisu.mec.gov.br/>> Acesso em: 01 fev. 2014

Sisutec - **Página Inicial**. Disponível em: <<http://sisutec.mec.gov.br/>> Acesso em: 01 fev. 2014

Wikipedia. **Allan Birnbaum**. Disponível em: < [http://en.wikipedia.org/wiki/Allan\\_Birnbaum](http://en.wikipedia.org/wiki/Allan_Birnbaum) > Acesso em: 01 fev. 2014

Wikipedia. **Frederic Lord**. Disponível em: < [http://en.wikipedia.org/wiki/Frederic\\_M.\\_Lord](http://en.wikipedia.org/wiki/Frederic_M._Lord) > Acesso em: 01 fev. 2014

**APÊNDICE**

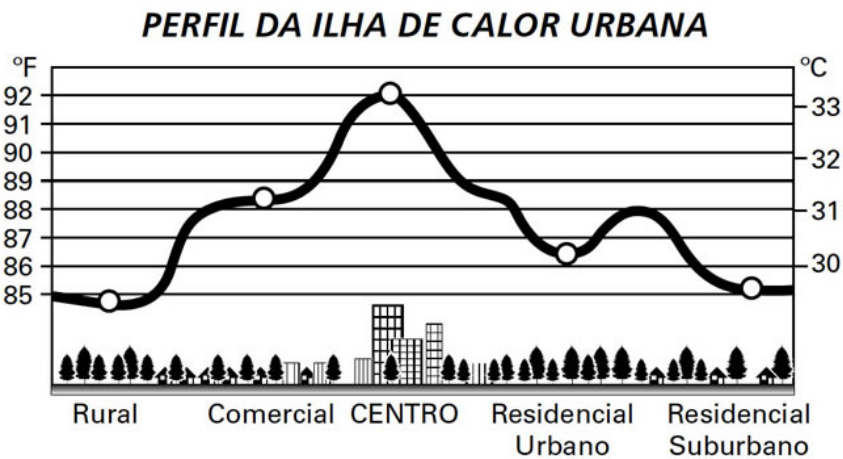
**APÊNDICE A - Pré-teste aplicado aos 216 respondentes, sobre o tema Tratamento da Informação (já excluída a questão que foi anulada).**

**QUESTIONÁRIO PARA AVALIAÇÃO DE CONHECIMENTOS EM MATEMÁTICA**

**TEMA: TRATAMENTO DA INFORMAÇÃO**

**CÓDIGO DO PARTICIPANTE: .....**

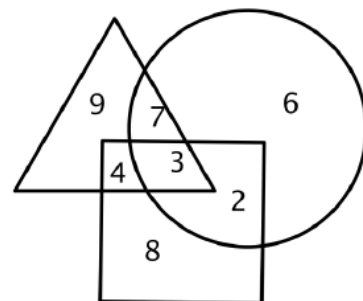
1) (Enem 2011) Rafael mora no Centro de uma cidade e decidiu se mudar, por recomendações médicas, para uma das regiões: Rural, Comercial, Residencial Urbano ou Residencial Suburbano. A principal recomendação médica foi com as temperaturas das “ilhas de calor” da região, que deveriam ser inferiores a  $31^{\circ}\text{C}$ . Tais temperaturas são apresentadas no gráfico:



Escolhendo, aleatoriamente, uma das outras regiões para morar, a probabilidade de ele escolher uma região que seja adequada às recomendações médicas é

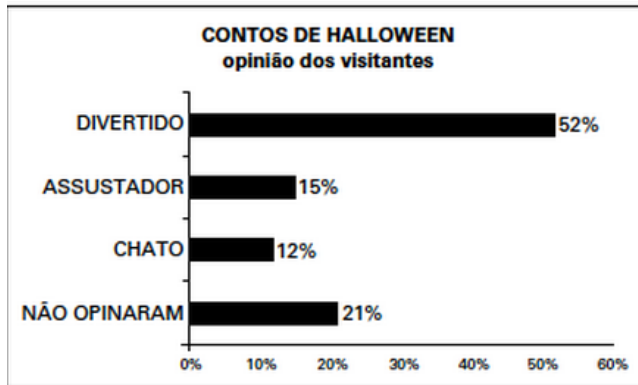
- a)  $\frac{1}{5}$                       b)  $\frac{1}{4}$                       c)  $\frac{2}{5}$                       d)  $\frac{3}{5}$                       e)  $\frac{3}{4}$

2) (UNICAMP 2013 - Adaptada) O diagrama ao lado indica a distribuição de alunos matriculados em três cursos de uma escola. Com base nas informações do diagrama, a probabilidade de um aluno escolhido ao acaso estar matriculado em apenas um curso é:



- a)  $\frac{9}{39}$                       d)  $\frac{16}{39}$   
 b)  $\frac{39}{6}$                       e)  $\frac{23}{39}$   
 c)  $\frac{39}{23}$

3) (ENEM 2012) Em um blog de variedades, músicas, mantras e informações diversas, foram postados “Contos de Halloween”. Após a leitura, os visitantes poderiam opinar, assinalando suas relações em: “Divertido”, “Assustador” ou “Chato”. Ao final de uma semana, o blog registrou que 500 visitantes distintos acessaram esta postagem. O gráfico a seguir apresenta o resultado da enquete.

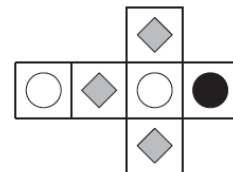


O administrador do blog irá sortear um livro entre os visitantes que opinaram na postagem “Contos de Halloween”. Sabendo que nenhum visitante votou mais de uma vez, a probabilidade de uma pessoa escolhida ao acaso entre as que opinaram ter assinalado que o conto “Contos de Halloween” é “Chato” é mais aproximada por

- a) 0,09.    b) 0,12.    c) 0,14.    d) 0,15.    e) 0,18.

4) (OBMEP 2013) Um dado foi construído usando a planificação da figura. Qual é a probabilidade de obtermos dois resultados diferentes quando jogamos esse dado duas vezes?

- a)  $\frac{1}{2}$                       c)  $\frac{2}{3}$                       e)  $\frac{31}{36}$   
 b)  $\frac{11}{18}$                       d)  $\frac{5}{6}$



5) (FUVEST 2011) Um dado cúbico, não viciado, com faces numeradas de 1 a 6, é lançado três vezes. Em cada lançamento, anota-se o número obtido na face superior do dado, formando-se uma sequência (a, b, c). Qual é a probabilidade de que b seja sucessor de a ou que c seja sucessor de b?

- a)  $\frac{4}{27}$                       b)  $\frac{11}{54}$                       c)  $\frac{7}{27}$                       d)  $\frac{10}{27}$                       e)  $\frac{23}{54}$

6) (UFF 2005) Niterói é uma excelente opção para quem gosta de fazer turismo

ecológico. Segundo dados da prefeitura, a cidade possui oito pontos turísticos dessa natureza. Certo hotel da região oferece de brinde a cada hóspede a possibilidade de escolher três dos oito pontos turísticos ecológicos para visitar durante sua estada. O número de modos diferentes com que um hóspede pode escolher, aleatoriamente, três destes locais, independentemente da ordem escolhida, é:

- a) 8                      b) 24                      c) 56                      d) 112                      e) 336

7) (Guia do Estudante) As antigas placas para automóveis, com duas letras seguidas de quatro algarismos, foram substituídas por novas com três letras seguidas de quatro algarismos. Nestas placas, bem como nas antigas, são utilizadas as 23 letras do alfabeto português, mais as letras K, W, Y. Quantos carros a mais puderam ser emplacados com o novo sistema?

- a)  $17576 \cdot 10^4$                       c)  $676 \cdot 10^5$                       e)  $169 \cdot 10^6$   
b)  $17576 \cdot 10^5$                       d)  $676 \cdot 10^6$

8) (UFAM 2006) Uma prova de matemática consta 8 questões das quais o aluno deve escolher 6. De quantas formas ele poderá escolher as 6 questões?

- a) 8                      b) 56                      c) 336                      d) 1680                      e) 28

9) (UNESP 2002) Quatro amigos, Pedro, Luísa, João e Rita, vão ao cinema, sentando-se em lugares consecutivos na mesma fila. O número de maneiras que os quatro podem ficar dispostos de forma que Pedro e Luísa fiquem sempre juntos e João e Rita fiquem sempre juntos é

- a) 2.                      b) 4.                      c) 8.                      d) 16.                      e) 24.

10) (FUVEST 2007) Em uma classe de 9 alunos, todos se dão bem, com exceção de Andréia, que vive brigando com Manoel e Alberto. Nessa classe, será constituída uma comissão de cinco alunos, com a exigência de que cada membro se relacione bem com todos os outros. Quantas comissões podem ser formadas?

- a) 71                      b) 75                      c) 80                      d) 83                      e) 87

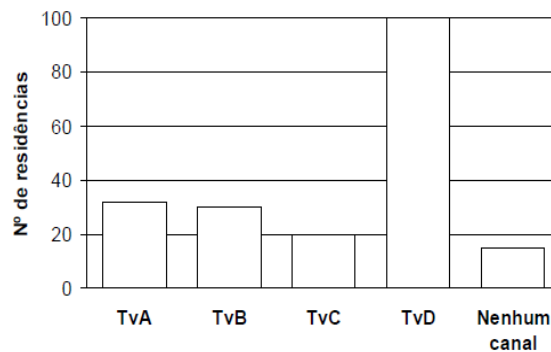
11) (FGV 2008 ADAPTADO) Sendo  $x$ ,  $y$  e  $z$  três números naturais e, diferentes de 1, tais que  $x \cdot y \cdot z = 2310$ , o número de conjuntos  $\{x, y, z\}$  diferentes é

- a) 32                      b) 36                      c) 40                      d) 43                      e) 45

12) (ENEM 98) Uma pesquisa de opinião foi realizada para avaliar os níveis de audiência de alguns canais de televisão, entre 20h e 21h, durante uma determinada noite. Os resultados obtidos estão representados no gráfico de barras abaixo:

O número de residências atingidas nessa pesquisa foi **aproximadamente** de:

- a) 100                      d) 200  
b) 135                      e) 220  
c) 150



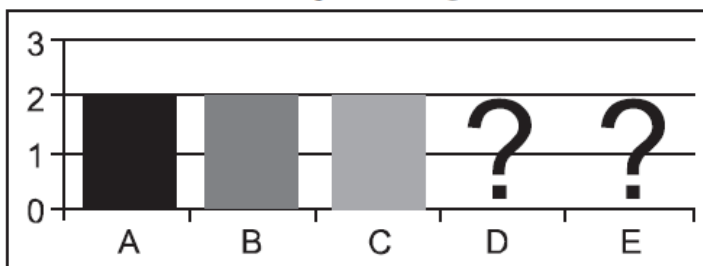
13) (FGV 2008) Sejam os números

7, 8, 3, 5, 9 e 5 seis números de uma lista de nove números inteiros. O maior valor possível para a mediana dos nove números da lista é

- a) 5                      b) 6                      c) 7                      d) 8                      e) 9

14) (ENEM 2009 cancelado) Cinco equipes A, B, C, D e E disputaram uma prova de gincana na qual as pontuações recebidas podiam ser 0, 1, 2 ou 3. A média das cinco equipes foi de 2 pontos. As notas das equipes foram colocadas no gráfico a seguir, entretanto, esqueceu-se de representar as notas da equipe D e da equipe E. Mesmo sem aparecer as notas das equipes D e E, pode-se concluir que os valores da moda e da mediana são, respectivamente,

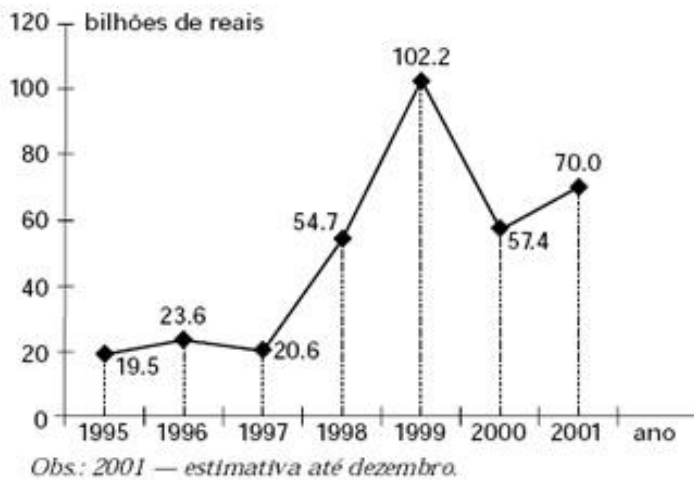
**Pontuação da gincana**



- a) 1,5 e 2,0.  
b) 2,0 e 1,5.  
c) 2,0 e 2,0.  
d) 2,0 e 3,0.  
e) 3,0 e 2,0.

15) (UNESP 2002) O gráfico, publicado na Folha de S. Paulo de 16.08.2001, mostra os gastos (em bilhões de reais) do governo federal com os juros da dívida pública.





Pela análise do gráfico, pode-se afirmar que:

- Em 1998, o gasto foi de R\$ 102,2 bilhões.
- O menor gasto foi em 1996.
- Em 1997, houve redução de 20% nos gastos, em relação a 1996.
- A média dos gastos nos anos de 1999 e 2000 foi de R\$ 79,8 bilhões.
- Os gastos decresceram de 1997 a 1999.

16) (ENEM 2010) Marco e Paulo foram classificados em um concurso. Para classificação no concurso o candidato deveria obter média aritmética na pontuação igual ou superior a 14. Em caso de empate na média, o desempate seria em favor da pontuação mais regular. No quadro a seguir são apresentados os pontos obtidos nas provas de Matemática, Português e Conhecimentos Gerais, a média, a mediana e o desvio padrão dos dois candidatos. Dados dos candidatos no concurso

	Matemática	Português	Conhecimentos Gerais	Média	Mediana	Desvio Padrão
Marco	14	15	16	15	15	0,32
Paulo	8	19	18	15	18	4,97

O candidato com pontuação mais regular, portanto mais bem classificado no concurso é:

- Marco, pois a média e a mediana são iguais.
- Marco, pois obteve menor desvio padrão.
- Paulo, pois obteve a maior pontuação da tabela, 19 em Português.

- d) Paulo, pois obteve maior mediana.
- e) Paulo, pois obteve maior desvio padrão.

17) (UFU 2008) Um concurso avaliou  $n$  candidatos atribuindo-lhes notas de 0 a 100 pontos. Sabe-se que exatamente 20 deles obtiveram nota máxima e, nesse caso, a média aritmética foi de 80 pontos. Agora, se consideradas apenas as notas inferiores a 100 pontos, a média passa a ser de 70 pontos. Nessas condições, pode-se afirmar que  $n$  é igual a

- a) 70
- b) 60
- c) 80
- d) 40
- e) 50

**APÊNDICE B - Tabela com os vetores de resposta do pré-teste aplicado aos 216 respondentes.**

		<b>ITENS</b>																
		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>
<b>1</b>		0	1	1	1	0	0	1	0	1	0	0	1	0	0	0	1	1
<b>2</b>		1	1	0	1	0	1	1	0	0	1	0	1	0	0	1	0	1
<b>3</b>		1	1	0	1	0	0	0	1	0	1	0	0	0	0	1	0	0
<b>4</b>		0	1	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0
<b>5</b>		1	0	1	0	0	0	0	1	1	0	0	0	1	1	1	1	0
<b>6</b>		1	1	1	1	0	1	0	0	0	1	0	0	0	1	0	0	1
<b>7</b>		1	0	1	1	0	1	0	0	1	1	0	0	1	0	0	1	0
<b>8</b>		0	1	1	0	1	1	0	0	1	0	1	0	0	1	1	1	0
<b>9</b>		0	1	1	0	0	1	1	1	0	0	1	0	0	0	1	0	1
<b>10</b>		1	0	0	1	0	1	0	1	0	0	0	1	1	1	1	1	0
<b>11</b>		1	1	1	0	1	0	0	0	0	1	0	1	0	0	0	1	0
<b>12</b>		0	1	1	0	1	1	0	0	0	0	1	1	0	0	1	1	0
<b>13</b>		1	1	1	0	0	1	0	0	1	0	0	1	0	0	0	0	1
<b>14</b>		0	1	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0
<b>15</b>		0	1	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0
<b>16</b>		1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0
<b>17</b>		1	1	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0
<b>18</b>		1	0	1	1	0	1	0	1	0	0	0	1	0	0	0	0	1
<b>19</b>		0	1	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0
<b>20</b>		1	1	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0
<b>21</b>		1	1	1	1	0	0	1	0	0	1	1	0	1	0	0	0	1
<b>22</b>		1	1	0	0	0	0	0	1	1	0	0	1	0	0	1	1	0
<b>23</b>		1	0	1	0	0	0	0	1	0	1	0	1	1	0	0	0	1
<b>24</b>		0	0	1	0	0	1	0	0	1	1	1	1	0	0	1	1	0
<b>25</b>		0	0	0	0	1	1	0	1	1	0	0	0	0	1	1	0	1
<b>26</b>		0	1	1	1	0	1	0	0	1	1	0	0	1	1	0	0	0
<b>27</b>		1	1	0	0	1	0	0	0	1	1	1	0	0	1	1	1	0
<b>28</b>		1	0	1	0	1	1	1	1	0	0	1	0	0	0	1	1	0
<b>29</b>		0	1	1	0	0	1	0	1	0	1	0	0	0	0	0	1	0
<b>30</b>		1	0	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0
<b>31</b>		1	0	0	0	0	1	0	1	1	1	0	1	1	1	0	1	0
<b>32</b>		0	1	1	0	0	1	0	1	0	0	1	0	0	1	1	1	0
<b>33</b>		0	1	0	0	1	1	0	1	1	0	1	0	0	1	1	0	0

---

<b>34</b>	0	1	1	1	1	0	0	0	1	0	0	1	0	1	1	0	0
<b>35</b>	0	0	1	1	0	0	1	0	1	0	0	1	0	1	1	1	0
<b>36</b>	1	1	0	1	0	1	0	0	0	1	1	1	0	0	0	1	0
<b>37</b>	1	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0
<b>38</b>	0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	1	1
<b>39</b>	0	0	1	0	1	0	0	1	1	1	0	1	0	1	0	0	0
<b>40</b>	1	1	1	1	0	0	1	0	0	1	1	0	0	0	0	1	0
<b>41</b>	1	0	1	1	0	1	0	1	0	0	1	0	0	1	0	0	1
<b>42</b>	0	1	1	0	0	1	1	1	0	0	0	1	0	0	1	1	0
<b>43</b>	1	0	0	1	0	0	0	1	1	0	0	1	1	0	0	1	0
<b>44</b>	1	1	0	1	0	1	1	0	0	0	0	0	1	0	0	1	1
<b>45</b>	0	1	1	0	0	1	0	0	1	0	1	0	0	1	0	1	0
<b>46</b>	1	0	1	0	1	0	0	1	1	1	0	1	1	0	0	1	0
<b>47</b>	1	0	0	1	1	0	0	0	1	1	0	1	0	0	1	0	1
<b>48</b>	1	0	0	1	1	1	0	0	1	1	0	1	1	0	0	1	0
<b>49</b>	1	1	1	1	0	1	1	1	0	1	0	1	0	1	1	0	0
<b>50</b>	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0
<b>51</b>	1	1	1	0	0	1	0	1	0	1	0	0	1	0	0	1	0
<b>52</b>	1	0	0	1	0	0	0	1	1	1	0	1	0	0	0	1	0
<b>53</b>	1	1	0	1	0	0	1	1	1	1	0	0	1	1	0	1	0
<b>54</b>	1	1	1	0	0	1	0	1	0	1	0	0	0	1	1	0	0
<b>55</b>	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0
<b>56</b>	0	1	1	0	0	1	0	0	1	0	0	1	1	0	0	0	1
<b>57</b>	1	0	0	1	1	0	1	0	0	0	1	0	0	0	0	1	0
<b>58</b>	1	0	1	0	0	1	0	1	1	0	0	1	0	0	0	1	0
<b>59</b>	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0
<b>60</b>	0	1	1	0	0	0	0	0	1	1	0	0	1	0	0	1	0
<b>61</b>	0	1	0	1	0	1	0	1	1	1	0	0	1	0	0	0	0
<b>62</b>	0	0	1	1	0	0	1	1	0	1	1	1	1	0	1	1	0
<b>63</b>	1	0	1	1	0	1	0	0	1	1	0	1	1	0	0	1	0
<b>64</b>	0	1	0	1	0	1	1	0	0	1	0	1	1	1	0	0	0
<b>65</b>	0	1	1	0	1	1	1	1	0	1	0	1	1	1	1	0	0
<b>66</b>	1	0	1	0	0	1	0	1	1	0	1	1	0	0	0	1	0
<b>67</b>	0	1	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0
<b>68</b>	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	0
<b>69</b>	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0
<b>70</b>	1	0	1	1	0	0	0	0	1	1	0	1	0	1	1	1	0

---

---

<b>71</b>	1	1	1	0	0	1	0	1	1	0	0	0	1	1	1	0	0
<b>72</b>	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	1	1
<b>73</b>	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
<b>74</b>	0	1	0	0	0	0	0	0	1	1	1	1	0	1	1	0	0
<b>75</b>	0	1	0	0	0	0	0	0	1	0	0	1	0	1	1	1	1
<b>76</b>	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
<b>77</b>	0	1	0	0	1	0	0	0	1	0	0	1	0	1	1	1	0
<b>78</b>	1	0	0	0	0	0	0	0	1	1	1	1	0	1	1	1	0
<b>79</b>	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
<b>80</b>	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
<b>81</b>	0	1	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0
<b>82</b>	0	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0
<b>83</b>	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
<b>84</b>	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	0
<b>85</b>	0	1	1	0	1	0	1	1	0	0	1	1	0	1	1	0	0
<b>86</b>	1	1	0	0	0	1	0	0	1	0	0	1	0	1	1	1	0
<b>87</b>	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1	1	0
<b>88</b>	1	1	0	1	0	0	1	1	0	0	0	1	1	0	0	1	1
<b>89</b>	0	1	1	1	0	1	1	0	1	0	0	1	0	1	1	0	1
<b>90</b>	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	0
<b>91</b>	1	1	0	1	1	0	1	0	1	0	0	0	0	1	0	1	0
<b>92</b>	1	1	0	0	0	1	1	0	1	0	0	1	1	1	0	1	0
<b>93</b>	0	1	1	0	0	1	0	0	0	0	0	0	0	1	1	1	0
<b>94</b>	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	0
<b>95</b>	1	1	0	1	0	1	0	0	0	0	0	1	0	1	1	0	1
<b>96</b>	1	1	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1
<b>97</b>	0	1	1	1	0	0	1	1	0	1	0	1	1	1	1	1	0
<b>98</b>	0	1	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0
<b>99</b>	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	1
<b>100</b>	1	1	0	1	0	0	1	1	0	0	0	1	0	1	0	1	1
<b>101</b>	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
<b>102</b>	0	1	1	1	0	0	0	1	0	1	0	1	1	0	1	0	0
<b>103</b>	0	1	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0
<b>104</b>	1	1	0	0	0	1	0	0	0	0	0	1	0	1	1	1	0
<b>105</b>	0	1	0	0	0	1	0	0	0	0	0	1	0	1	1	1	1
<b>106</b>	0	1	0	1	0	1	1	0	1	1	0	1	0	0	1	0	1
<b>107</b>	0	1	0	0	1	0	0	0	1	1	1	1	0	1	1	0	0

---

---

<b>108</b>	1	1	1	1	1	0	0	0	1	0	0	1	0	1	1	0	0
<b>109</b>	0	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0
<b>110</b>	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0
<b>111</b>	0	1	0	1	0	0	0	0	1	1	0	1	0	0	0	1	0
<b>112</b>	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
<b>113</b>	0	1	0	0	0	1	0	0	0	0	0	1	0	1	1	1	0
<b>114</b>	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0
<b>115</b>	0	1	1	0	0	0	1	0	1	0	0	1	0	0	0	1	0
<b>116</b>	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
<b>117</b>	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
<b>118</b>	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0
<b>119</b>	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
<b>120</b>	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1
<b>121</b>	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
<b>122</b>	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	1	0
<b>123</b>	0	0	0	1	1	1	0	0	0	0	0	1	0	1	0	1	1
<b>124</b>	1	0	0	0	0	1	0	0	1	0	0	1	0	0	1	1	0
<b>125</b>	0	1	0	1	1	0	0	0	0	0	0	1	0	1	1	1	0
<b>126</b>	0	0	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0
<b>127</b>	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0
<b>128</b>	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
<b>129</b>	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0
<b>130</b>	0	1	0	0	0	0	1	0	0	1	0	1	0	1	1	1	0
<b>131</b>	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
<b>132</b>	0	1	0	0	1	0	0	0	0	0	1	1	0	1	1	1	1
<b>133</b>	1	0	0	0	1	0	1	0	1	0	0	1	0	1	1	1	0
<b>134</b>	0	1	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0
<b>135</b>	0	0	0	1	0	0	1	0	0	1	0	1	0	1	1	0	0
<b>136</b>	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	1
<b>137</b>	0	1	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0
<b>138</b>	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0
<b>139</b>	0	1	0	0	0	1	1	1	0	0	1	1	0	1	0	0	0
<b>140</b>	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1
<b>141</b>	0	1	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0
<b>142</b>	0	1	0	0	1	1	0	0	0	0	0	1	0	1	1	0	0
<b>143</b>	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
<b>144</b>	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	0	0

---

---

<b>145</b>	0	1	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0
<b>146</b>	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
<b>147</b>	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0
<b>148</b>	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
<b>149</b>	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
<b>150</b>	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
<b>151</b>	0	0	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0
<b>152</b>	1	1	0	0	0	0	1	0	0	0	0	1	0	1	1	0	1
<b>153</b>	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
<b>154</b>	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
<b>155</b>	1	1	0	0	0	1	1	0	1	0	0	1	0	1	1	0	0
<b>156</b>	0	1	0	0	1	1	0	1	1	0	0	1	0	1	1	0	1
<b>157</b>	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
<b>158</b>	0	1	0	1	0	0	0	1	0	0	1	1	0	1	1	1	0
<b>159</b>	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
<b>160</b>	0	1	0	1	0	1	1	1	0	0	0	0	0	1	0	1	0
<b>161</b>	0	1	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0
<b>162</b>	0	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	0
<b>163</b>	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	1
<b>164</b>	0	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
<b>165</b>	1	1	0	0	0	1	0	1	0	1	0	0	0	0	1	1	0
<b>166</b>	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
<b>167</b>	0	0	1	0	0	1	0	1	0	0	1	1	0	1	1	1	0
<b>168</b>	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
<b>169</b>	0	1	0	0	1	0	0	1	0	0	1	0	1	1	0	1	0
<b>170</b>	0	1	0	0	1	0	0	1	0	1	0	1	0	1	1	0	0
<b>171</b>	0	1	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0
<b>172</b>	1	1	0	0	1	1	0	1	0	0	0	0	0	1	1	1	0
<b>173</b>	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
<b>174</b>	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
<b>175</b>	0	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0
<b>176</b>	1	1	0	0	0	1	0	1	0	0	1	0	1	0	1	0	0
<b>177</b>	1	1	0	1	0	0	1	1	1	0	0	0	1	1	1	1	0
<b>178</b>	1	1	0	0	0	0	1	1	1	0	1	1	1	1	1	1	0
<b>179</b>	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	1	0
<b>180</b>	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1
<b>181</b>	0	1	0	0	0	0	0	1	0	1	0	1	0	1	1	0	1

---

---

<b>182</b>	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0
<b>183</b>	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
<b>184</b>	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0
<b>185</b>	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
<b>186</b>	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0
<b>187</b>	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0
<b>188</b>	0	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	0
<b>189</b>	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	1	0
<b>190</b>	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
<b>191</b>	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0	1	0
<b>192</b>	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1
<b>193</b>	0	1	0	1	0	1	0	0	1	0	0	1	0	1	0	0	0
<b>194</b>	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0
<b>195</b>	0	1	0	0	0	0	0	0	0	1	0	1	0	1	1	1	0
<b>196</b>	0	1	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0
<b>197</b>	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0
<b>198</b>	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0
<b>199</b>	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
<b>200</b>	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
<b>201</b>	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1
<b>202</b>	1	1	0	0	0	1	0	0	0	0	1	1	0	1	0	1	0
<b>203</b>	1	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0
<b>204</b>	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
<b>205</b>	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0	0	0
<b>206</b>	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
<b>207</b>	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0
<b>208</b>	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	1	0
<b>209</b>	1	1	1	0	1	0	0	1	1	1	0	1	0	1	1	1	1
<b>210</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
<b>211</b>	0	1	0	0	0	1	1	0	0	0	0	1	0	1	1	0	0
<b>212</b>	0	1	0	0	0	0	0	0	1	0	0	1	0	1	1	1	0
<b>213</b>	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0
<b>214</b>	0	1	0	0	0	1	0	1	1	1	0	0	0	1	1	0	1
<b>215</b>	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
<b>216</b>	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0

---



## APÊNDICE C - Rotina criado no Bloco de Notas para implementação da análise dos dados da Tabela pelo R.

```

# Script para realizar uma análise de TRI
# Primeiro Mudar o diretório para a pasta onde está o arquivo
# Para mudar diretório: Clicar no console; arquivo;mudar diretorio;buscar a pasta

dados=read.table("Resultado_questionario.txt",header=TRUE)
dados
resultado.itens=dados[,2:18]
resultado.itens
library(ltm) # carregar pacote ltm

#####aqui só algumas coisas nos interessam: Teoria Clássica#####
resultado.desc=describe(resultado.itens)
resultado.desc

#####

library(irtoys)

resultado.par=est(resultado.itens,model="3PL",engine="ltm",nqp=10)
resultado.par

resultado.par2=resultado.par$est
resultado.par2
round(resultado.par2,4)

erro_padrao=resultado.par$se
round(erro_padrao,4)

resultado.sco=eap(resultado.itens,resultado.par2,qu=normal.qu())
habilidade_logistica=round(resultado.sco[,1],4) # Resultado da estimativa da habilidade na escala logistica
habilidade_logistica
|
# Mudança de Escala

library(CTT)
novo.score=score.transform(resultado.sco[,1], mu.new=500,sd.new=100,normalize=FALSE)
round(head(novo.score$new,n=216),2)

sd(resultado.sco[,1])

```