



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS  
PROGRAMA DE PÓS-GRADUAÇÃO  
EM MATEMÁTICA E ESTATÍSTICA

Rosa Maria Lobo Rosário

Interpretação de Escala em Avaliação Educacional: O Caso do  
Núcleo Pedagógico Integrado (NPI)

Orientador: Prof. Dr. Héilton Ribeiro Tavares

Belém  
2008

**Rosa Maria Lobo Rosário**

**Interpretação de Escala em Avaliação Educacional: O Caso do  
Núcleo Pedagógico Integrado (NPI)**

Dissertação de Conclusão de Curso apresentada ao Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará para obtenção do Título de Mestre em Estatística.

Orientador: **Prof. Dr. Héilton Ribeiro Tavares**

Belém  
2008

**Rosa Maria Lobo Rosário**

**Interpretação de Escala em Avaliação Educacional: O Caso do  
Núcleo Pedagógico Integrado (NPI)**

Dissertação apresentada para a obtenção do grau de Mestre em Estatística, Programa de Pós-Graduação em Matemática e Estatística, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará.

**Belém, 13 de junho de 2008**

---

Prof. *Dr.* Mauro de Lima Santos  
(Coordenador do Programa de Pós-Graduação em Matemática e Estatística)

**Banca Examinadora**

---

Prof. *Dr.* Héilton Ribeiro Tavares  
UFPA/INEP - **Orientador**

---

Prof. *Dr.* Amaury Patrick Gremaud  
USP/INEP - **Membro Externo**

---

Prof. *Dr.* Marcus Pinto da Costa da Rocha  
UFPA - **Membro Interno**

---

Prof. *Dr.* Valcir João da Cunha Farias  
UFPA - **Membro Interno**

*DEDICATÓRIA*

*Aos Meus Pais Alfredo de Barros Lobo (in memoriam)  
e Francisca da Costa Lobo (in memoriam)*

*À Minha Avó Laurinda Barbosa da Costa (in memoriam)*

*Ao meu Esposo Irandir da Silva Rosário e*

*Aos meus Filhos: Eduardo Lobo Rosário  
e Rosana Lobo Rosário*

*“meus tesouros”*

## Agradecimentos

À Deus que concedeu-me a vida, tranqüiliza-me espiritualmente e que deu-me a capacidade de realizar este trabalho.

Aos meus pais (in memoriam) Alfredo de Barros Lobo e Francisca da Costa Lobo, e à minha avó (in memoriam) Laurinda Barbosa da Costa, que foram os meus primeiros mestres, meus orientadores no caminho certo da vida, sempre apoiaram minhas decisões e deram o suporte necessário para concretizar meus objetivos.

Ao meu esposo Irandir Rosário e aos meus filhos Eduardo e Rosana, pelas milhares de horas que abdiquei de suas companhias, e aos meus irmãos: Ana Maria, Luzia Maria, Francisco Cezar, Elza Maria e Manoel Evilázio que sempre acreditaram em mim.

Ao meu orientador, Prof. Dr. Héilton Ribeiro Tavares, integrante do Programa de Pós-Graduação em Matemática e Estatística, pela orientação competente, dedicando-se de forma marcante na elaboração deste trabalho.

À Universidade Federal do Pará, de modo especial ao Núcleo Pedagógico Integrado e a todos que lá trabalham, destacando-se os amigos: Luís Marconi, Édna Lima, José Farias, Walter, Danielle, Vera Lúcia, Vena, Izabel Cristina, Waldelice, Helena Lima, Neuza, Nazaré, Cleuma, Analice, Eunice, Eliete, Madalena e à querida amiga Elza de Nazaré de Oliveira Nilander (in memoriam).

Ao professor Mauro de Lima Santos Rocha, coordenador do Programa de Pós-Graduação em Matemática e Estatística.

Aos alunos da 4<sup>a</sup> e da 8<sup>a</sup> séries de 2006, da Escola de Aplicação (NPI), que participaram da pesquisa.

Ao INEP, pelas informações necessárias à realização desta dissertação.

Ao Prof. MsC José Gracildo de C. Júnior, pela valiosa ajuda na utilização do “Latex”.

À todos os professores integrantes do Programa de Pós-Graduação em Matemática e Estatística, e aos professores do curso de graduação em Estatística, pela contribuição valiosa e significativa ao longo do Curso de Graduação e de Mestrado.

Ao diretor do DAVES, Prof. Dr. Luís Acácio Cordeiro Centeno e aos funcionários: Carlos Alexandre Silva e Sebastião de Lima Serdeira que imprimiram as provas e a Noraney Bandeira Alves que imprimiu os questionários.

Aos colegas e amigos: Adriana Carvalho e Natália Barros do grupo de estudo, Risoleide Barbosa, Elisangela Brasil, Giovana Rodrigues, Luciane Barbosa, Sheyla Barbosa, que colaboraram na aplicação das provas, Leandro, Jaciane, Marco Pollo, Paulo Fernando, Edilson Torres, Félix Lelis, Michelle Borges, Renatão, Ana Cristina, Suzy, Suellen Arruda, Reneide Souza, Marta Luzilena Silva e finalmente agradeço às todos que de alguma forma contribuíram para a realização desta dissertação.

---

## Resumo

---

A Teoria da Resposta ao Item (TRI) é uma técnica que vem sendo largamente empregada em avaliação educacional, principalmente porque ela permite uma construção de escala de proficiências que é impraticável quando usa-se apenas os escores individuais, bem como porque ela produz, teoricamente, resultados independentes do conjunto de itens adotado. Esta técnica já é aplicada nas principais avaliações em larga escala realizadas no país e nas principais avaliações internacionais. Neste trabalho estuda-se a TRI, em seus aspectos teóricos e computacionais, bem como faz-se uma aplicação de testes de Língua Portuguesa e Matemática, na Escola de Aplicação da UFPA (Núcleo Pedagógico Integrado - NPI), culminando na alocação do NPI na escala única de proficiência nacional, seja de Língua Portuguesa ou Matemática.

Palavras-Chave: avaliação educacional, TRI, escala de proficiência.

---

## Abstract

---

The Item Response Theory (IRT) is a technique that is being widely used in educational evaluation, principally because it allows a construction of a scale of proficiencies that is impracticable when it is used just the individual scores, and it produces item independent score results. This technique is applied in the main evaluations in wide scale that are realized in the country, as well as the main international evaluations. In this paper it is studied the IRT, in its theoretical and computational aspects, as well as the realization of application of some Portuguese Language and mathematics Tests at Escola de Aplicação da UFPA (Núcleo Pedagógico Integrado - NPI), culminating in the allocation of NPI in the only scale of the national proficiency, as well in Portuguese Language as in Mathematics.

Keywords: educational evaluation, IRT, knowledge scales.

---

# Conteúdo

---

<b>Resumo</b>	<b>vi</b>
<b>Abstract</b>	<b>vii</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>Lista de Figuras</b>	<b>xii</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
1.1 Avaliação Educacional . . . . .	1
1.1.1 Objetivo do Trabalho . . . . .	3
1.1.2 Importância e Justificativa do Trabalho . . . . .	3
<b>2 SISTEMAS DE AVALIAÇÃO REALIZADOS NO BRASIL</b>	<b>4</b>
2.1 Introdução . . . . .	4
2.2 Programa Internacional de Avaliação de Alunos (Pisa) . . . . .	5
2.3 Exame Nacional do Ensino Médio (Enem) . . . . .	9
2.4 Exame Nacional de Certificação de Competências para o Ensino de Jovens e Adultos (Encceja) . . . . .	11
2.4.1 Estrutura do Encceja . . . . .	12
2.4.2 Matriz de Competência e Habilidades do Encceja . . . . .	12
2.5 Sistema de Avaliação da Educação Básica (Saeb) . . . . .	13
<b>3 SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA (Saeb)</b>	<b>14</b>
3.1 Introdução . . . . .	14
3.2 Histórico . . . . .	14
3.3 Metodologias do Saeb . . . . .	17
3.3.1 Matrizes de Referência . . . . .	17
3.3.2 Testes Padronizados . . . . .	18
3.3.3 Bloco Incompleto Balanceado (BIB) . . . . .	19
3.3.4 Escalas de Proficiência . . . . .	20
3.3.5 Teoria da Resposta ao Item . . . . .	20
3.4 Resultados do Saeb . . . . .	20
3.5 A Prova Brasil . . . . .	21
3.6 O Ideb . . . . .	21

---

<b>4</b>	<b>TEORIA CLÁSSICA DOS TESTES</b>	<b>23</b>
4.1	Introdução . . . . .	23
4.2	O Modelo da Psicometria Clássica: TCT . . . . .	24
4.3	Análise Algébrica dos Itens . . . . .	24
<b>5</b>	<b>TEORIA DA RESPOSTA AO ITEM</b>	<b>27</b>
5.1	Introdução . . . . .	27
5.2	Modelos Matemáticos da Teoria Resposta ao Item . . . . .	28
5.2.1	Modelo Logístico de Um Parâmetro (ML1) . . . . .	32
5.2.2	Modelo Logístico de Dois Parâmetros (ML2) . . . . .	34
5.2.3	Modelo Logístico de Três Parâmetros (ML3) . . . . .	35
5.3	Modelos de Múltiplos Grupos . . . . .	38
5.4	Métodos de Estimação . . . . .	39
5.4.1	Caso 1: já são conhecidas as habilidades dos respondentes e deseja-se estimar os parâmetros dos itens. . . . .	40
5.4.2	Caso 2: já são conhecidos os parâmetros dos itens e procura-se estimar as habilidades dos respondentes . . . . .	41
5.4.3	Caso 3: deseja-se estimar, simultaneamente, os parâmetros dos itens e as habilidades dos respondentes. . . . .	43
5.4.4	Máxima verossimilhança marginal . . . . .	43
5.5	Estimação das proficiências . . . . .	47
5.5.1	Estimação das Habilidades pela Média da Posteriori - EAP . . . . .	47
5.6	Sobre a escala . . . . .	48
5.7	Aspectos computacionais . . . . .	49
5.7.1	Principais softwares . . . . .	49
5.7.2	Bilog, Bilog-Mg . . . . .	50
<b>6</b>	<b>ESCALAS DE DESEMPENHO</b>	<b>52</b>
6.1	Introdução . . . . .	52
6.2	A Escala do Saeb . . . . .	53
<b>7</b>	<b>Resultado da Pesquisa: Posicionamento do NPI Escala do Saeb</b>	<b>55</b>
7.1	Introdução . . . . .	55
7.2	Resultado da Pesquisa . . . . .	55
7.2.1	Médias Gerais de Proficiência . . . . .	55
7.2.2	Desempenho dos Alunos por Estágios de Competências . . . . .	61
7.3	Conclusões . . . . .	62
<b>8</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>64</b>
	<b>Referências Bibliográficas</b>	<b>66</b>
<b>A</b>	<b>Matriz de Referência para Língua Portuguesa e Matemática</b>	<b>71</b>
<b>B</b>	<b>Brasil - Proficiências do Saeb 1995 - 2005</b>	<b>73</b>

C Escala de Desempenho - Matemática, 4 <sup>a</sup> série	74
D Escala de Desempenho - Matemática, 8 <sup>a</sup> série	78
E Escala de Desempenho - Língua Portuguesa, 4 <sup>a</sup> série	81
F Escala de Desempenho - Língua Portuguesa, 8 <sup>a</sup> série	83
G Método da Quadratura	85

---

## Lista de Tabelas

---

2.1	Distribuição dos alunos por série . . . . .	7
2.2	Participação e Proficiências Médias - Brasil . . . . .	8
3.1	Bloco Incompleto Balanceado - BIB . . . . .	19
5.1	Tabela com quatro itens com seus respectivos parâmetros . . . . .	37
7.1	Proficiências do Saeb - 2005 - (Brasil, Região Norte, Estado do Pará), e do NPI - 2006 . . . . .	58
B.1	Média de Proficiências em Língua Portuguesa . . . . .	73
B.2	Média de Proficiências em Matemática . . . . .	73

---

## Lista de Figuras

---

2.1	Pisa 2003 - Distribuição das escolas por região. Fonte: INEP . . . . .	7
2.2	Resultado do Pisa no Brasil, nos anos de 2000, 2003 e 2006. Fonte: INEP . . .	9
2.3	Associação entre Competências e Habilidades da Matriz do Enem . . . . .	10
5.1	Logística com $D=1,7$ vs Normal . . . . .	30
5.2	Diferença: Normal-logística . . . . .	31
5.3	Curva Característica de dois Itens típicos do ML1 . . . . .	33
5.4	Curva Característica de dois Itens típicos do ML2 . . . . .	34
5.5	Curva Característica de dois Itens típicos do ML3 . . . . .	36
5.6	CCI e FII dos Itens 1, 2, 3 e 4 (da Tabela 5.1) . . . . .	38
7.1	Brasil - Proficiências em Matemática (Saeb: 1995 - 2005 e NPI: 2006) . . .	57
7.2	Brasil - Proficiências em Língua Portuguesa (Saeb: 1995 - 2005 e NPI - 2006)	58
7.3	Proficiências Médias em Língua Portuguesa - 4ª série do Ensino Fundamen- tal do NPI, do Brasil, da Região Norte e do Estado do Pará. . . . .	59
7.4	Proficiências Médias em Matemática - 4ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará. . . . .	59
7.5	Proficiências Médias em Língua Portuguesa - 8ª série do Ensino Fundamen- tal do NPI, do Brasil, da Região Norte e do Estado do Pará. . . . .	60
7.6	Proficiências Médias em Matemática - 8ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará. . . . .	60
7.7	Percentual de alunos por Estágios de Competências - NPI 2006 . . . . .	62
G.1	. . . . .	85

# INTRODUÇÃO

---

## 1.1 Avaliação Educacional

Existe uma relação de proporcionalidade direta entre *Desenvolvimento de um País* e *Educação*, pois quanto maior for o investimento de um país em educação, maior será seu desenvolvimento. Através da avaliação educacional os governos tomam conhecimento de fatos como, por exemplo, se os recursos públicos aplicados em políticas educacionais estão propiciando uma escolarização de qualidade. No Brasil, a avaliação que antes era centrada no processo ensino-aprendizagem, generalizou-se para as mais diversas formas de avaliação institucional, como avaliação de instituições, de sistemas e de projetos ou políticas públicas. Essas avaliações têm sido usadas com os mais diversos propósitos. Hoje os recursos são distribuídos de acordo com metas estabelecidas pelas avaliações educacionais através de indicadores. O principal deles é o Índice de Desenvolvimento da Educação Básica (Ideb), proposto pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep). O Ideb é um indicador de qualidade educacional que combina informações de desempenho em exames padronizados (Prova Brasil ou Saeb) obtido pelos estudantes ao final das etapas de ensino (4<sup>a</sup> e 8<sup>a</sup> séries do ensino fundamental e 3<sup>a</sup> série do ensino médio) com informações sobre rendimento escolar (aprovação).

Com base no Ideb o Ministério da Educação lançou o Programa de Desenvolvimento da Educação (PDE), fixando metas e vinculando a liberação de recursos financeiros ao alcance dessas metas. O objetivo do MEC é provocar saltos qualitativos de forma que no ano de 2021 o resultado médio do Brasil seja equivalente ao resultado médio do Programa Internacional de Avaliação de Alunos (Pisa), um exame de caráter internacional composto por 30 países que formam a Organização para Cooperação e Desenvolvimento da Educação (OCDE). Nas três edições desta avaliação o Brasil participou como país convidado.

A demanda por educação tem aumentado a cada ano no país. O SAEB tem sido aplicado

desde 1990 com o objetivo de avaliar o sistema como um todo, de forma que sempre foi suficiente envolver uma amostra no estudo. Participam deste estudo alunos da rede pública municipal, estadual e federal, da área urbana e rural, e também a rede particular. Pelo aspecto amostral, sendo observadas apenas algumas escolas e, no máximo duas turmas das escolas selecionadas, não havia condições de apresentar resultados por escolas, mas apenas para um conjunto de 400 sub-estratos definidos pelos estratos citados acima, ou outros criados com base em variáveis de interesse, tais como participantes de programas sociais, por exemplo.

Com base nisso, surgiu uma forte pressão por parte das escolas e da sociedade para que os resultados fossem apresentados por escolas, assim criando uma extensão do Saeb conhecida hoje como Prova Brasil, pelo seu aspecto universal, que avalia todos os alunos da rede pública. Contudo, dada a dimensão do estudo acabou-se criando limitações, de forma que ele não envolve escolas da área rural e nem a rede particular.

Por conta dos resultados de baixo nível obtidos para a 4ª série, tanto em Língua Portuguesa quanto em Matemática, o MEC planeja identificar a origem das deficiências e orientar os professores sobre o que seus alunos dominam, e o que deve ser abordado para eliminar as deficiências. Para isso, será feita uma aplicação aos alunos da primeira série (segundo ano) de estudo. Esta avaliação está sendo denominada de Provinha Brasil, com aplicação prevista para o primeiro semestre de 2008. Uma avaliação similar no segundo semestre avaliará o quanto de conhecimento foi agregado aos estudantes.

Relativamente à estrutura desta dissertação, no Capítulo 2 apresenta-se com maiores detalhes os principais Sistemas de Avaliação Educacional aplicados no Brasil: o Programa Internacional de Avaliação de Alunos (PISA), o Exame Nacional do Ensino Médio (ENEM), o Exame Nacional de Competências para o Ensino de Jovens e Adultos (ENCCEJA), o Sistema de Avaliação da Educação Básica (SAEB/ANEB) e a Prova Brasil (ANRESC). No Capítulo 3 é apresentada a estrutura do Saeb. No Capítulo 4 é feita uma revisão da literatura sobre a Teoria Clássica dos Testes. No Capítulo 5 é feita uma revisão da literatura sobre a Teoria da Resposta ao Item (TRI). No Capítulo 6 estuda-se a Interpretação de Escalas. O Capítulo 7 mostra o resultado do estudo sobre a colocação do NPI (Escola de Aplicação da Universidade Federal do Pará) na escala do Saeb. No Capítulo 8 estão

as Considerações Gerais e por último são apresentadas as Referências Bibliográficas e o Apêndice.

### **1.1.1 Objetivo do Trabalho**

Estudar primeiramente a TRI em seus aspectos teóricos e computacionais através de uma aplicação de testes de Língua Portuguesa e Matemática aos alunos das 4<sup>a</sup> e 8<sup>a</sup> séries do Ensino Fundamental do NPI e depois posicionar essa escola na escala única de proficiência nacional.

### **1.1.2 Importância e Justificativa do Trabalho**

A importância deste trabalho está no fato de dominar os aspectos teóricos e computacionais da TRI bem como a possibilidade de colocar os alunos em uma escala comum, mesmo que eles não respondam a todos os itens. Desse modo é possível conhecer exatamente as deficiências de cada aluno e poder ajudá-los a superá-las.

---

## Capítulo 2

# SISTEMAS DE AVALIAÇÃO REALIZADOS NO BRASIL

---

### 2.1 Introdução

Etimologicamente, avaliar significa “determinar a valia ou o valor de”, e avaliação, “valor determinado pelos avaliadores” (Holanda, 1986, p.205).

A avaliação, que a princípio era realizada de maneira não sistemática, foi aos poucos adquirindo uma forma mais estruturada, tornando-se um subsídio indispensável no monitoramento das reformas e das políticas educacionais. Para Bloom et al. (1975), a avaliação pode ser considerada como um método de adquirir e processar evidências necessárias para melhorar o ensino e a aprendizagem, incluindo uma grande variedade de evidências que vão além do exame usual de “papel e lápis”.

Luckesi (1996) analisou a avaliação dentro de dois modelos: o Conservador (que se sustenta em uma prática escolar autoritária, que objetiva o controle e o exercício do poder pelo professor) e o Transformador (surge como forma de superação do autoritarismo e como estabelecimento da autonomia). No primeiro caso, a avaliação educacional é um instrumento disciplinador das condutas cognitivas e sociais, dentro do contexto escolar. No segundo caso ela é tida como um instrumento de diagnóstico, objetivando o avanço e o crescimento, e não a estagnação disciplinadora. Para ele, avaliar significa “diagnosticar e intervir, o que quer dizer praticar a investigação sobre o que esta acontecendo, tendo em vista proceder intervenções adequadas, sempre para a melhoria dos resultados” (Luckesi, 2005, p.8).

A Avaliação de Sistemas é realizada em redes de ensino e tem por objetivo gerar informações importantes sobre os problemas da aprendizagem e as medidas que dizem respeito à melhoria da qualidade do ensino. O levantamento e a aferição do desempenho do aluno é feito através de testes de conhecimento e de questionários específicos para dire-

tores e professores. São levantados dados que possibilitam estimar os valores dos recursos recebidos e arrecadados pela escola.

As avaliações de sistemas ou avaliações em larga escala têm o objetivo de analisar a qualidade da educação baseada nos resultados de desempenho dos alunos de escolas públicas e privadas. São avaliados um número bastante grande de estudantes e os critérios utilizados são determinados por metodologia específica. O Sistema Nacional de Avaliação da Educação Básica (Saeb), o Exame Nacional de Competências para o Ensino de Jovens e Adultos (Encceja) e o Exame Nacional do Ensino Médio (Enem) são exemplos de avaliação em larga escala no âmbito nacional. O Programa Internacional de Avaliação de Alunos (Pisa) é uma avaliação em larga escala, de caráter internacional. A seguir apresenta-se uma breve descrição destas avaliações.

## **2.2 Programa Internacional de Avaliação de Alunos (Pisa)**

O Pisa (Programa Internacional de Avaliação de Alunos ou **P**rogramme for **I**nternational **S**tudent **A**ssessment) é uma avaliação internacional de habilidades e conhecimentos, desenvolvida pela OCDE (Organização para Cooperação e Desenvolvimento Econômico) para avaliar o desempenho de alunos na faixa dos 15 anos (idade em que se acredita que seja o término da escolaridade básica obrigatória na maioria dos países). É uma pesquisa em larga escala que testa internacionalmente conhecimentos e habilidades necessárias para a participação na sociedade e tem como principal objetivo produzir indicadores que contribuam para a discussão, dentro e fora dos países participantes, da qualidade da educação básica ministrada pelas escolas, para subsidiar políticas nacionais de melhoria da educação.

O Pisa tem como proposta, avaliar periodicamente (de três em três anos), um leque amplo de conhecimentos, habilidades e competências nas áreas de Leitura, Matemática e Ciências. Além de fazerem as provas, os alunos respondem um questionário de pesquisa socioeconômica e cultural.

Foi construída uma escala para cada uma das disciplinas avaliadas de forma que, no conjunto dos países da OCDE, em cada domínio, a média fosse de 500 pontos, com desvio padrão 100.

Do Pisa 2000 participaram 265.000 estudantes de 32 países sendo 28 pertencentes a OCDE. Neste ciclo foi avaliada a “habilidade dos alunos para resolver desafios do mundo real valendo-se de seu conhecimento e de suas competências” (OCDE, 2001). Os alunos fizeram provas de Literacia em Leitura\*, em Matemática† e em Ciências‡, sendo o foco maior em Leitura. O Brasil participou do Pisa 2000 como país convidado, com 4.893 estudantes. A média brasileira em Leitura foi de 383 pontos (último lugar), a média do conjunto dos países foi de 499 pontos. Em matemática, a média brasileira foi de 320 pontos e foi também o último lugar, a média do conjunto dos países foi de 498 pontos. Em ciências, a média brasileira foi de 365 pontos, novamente o último lugar e a média do conjunto dos países foi de 498 pontos.

No ano seguinte foi aplicado o “Pisa Ampliado” e a colocação do Brasil modificou, os estudantes brasileiros passaram a ocupar a 37ª posição na prova de Leitura, com uma nota média de 396, numa escala de zero a 800. Na Tabela 2.2 pode-se ver a pontuação dos brasileiros em Ciências, Leitura e Matemática nos ciclos de 2000, 2003 e 2006 (Inep).

Do segundo ciclo do Pisa, em 2003, participaram 250 mil adolescentes com 15 anos de idades em 41 países sendo 30 deles membros da OCDE. Da América Latina, participaram o Brasil, o Uruguai e o México. Os alunos fizeram três avaliações, sendo que o foco neste ciclo foi em Literacia Matemática.

\* Literacia em Leitura é definido como a capacidade de cada indivíduo compreender, usar textos escritos e refletir sobre eles, de modo a atingir os seus objetivos, a desenvolver os seus próprios conhecimentos e potencialidades e a participar ativamente na sociedade (OECD, 2001).

† A literacia matemática no Pisa é definida como a capacidade de um indivíduo identificar e compreender o papel que a matemática desempenha no mundo, de fazer julgamentos bem fundamentados e de usar e se envolver na resolução matemática das necessidades da sua vida, enquanto cidadão construtivo, preocupado e reflexivo (OCDE, 2001).

‡ A literacia em Ciências envolve o uso de conceitos científicos necessários para compreender e ajudar a tomar decisões sobre o mundo natural. Envolve também a capacidade de reconhecer questões científicas, fazer uso de evidências, tirar conclusões com bases científicas e comunicar essas conclusões (OCDE, 2001)

No Brasil participaram alunos de escolas públicas e privadas das cinco regiões (distribuídas entre estabelecimentos das zonas urbana e rural), num total de 229 escolas, em 179 municípios das cinco regiões. Fizeram parte da amostra alunos de 15 anos matriculados na 7<sup>a</sup> e 8<sup>a</sup> séries do Ensino Fundamental, ou na 1<sup>a</sup> ou 2<sup>a</sup> ou 3<sup>a</sup> do Ensino Médio. Foram selecionados cerca de 5.235 alunos distribuídos com base no Censo Escolar como mostra a Figura 2.1.

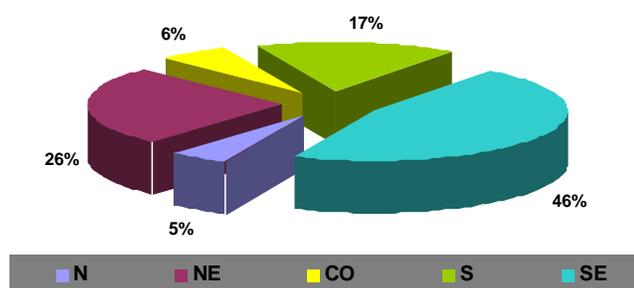


Figura 2.1 *Pisa 2003 - Distribuição das escolas por região. Fonte: INEP*

A distribuição dos alunos por série foi feita de acordo com a tabela abaixo:

Tabela 2.1 *Distribuição dos alunos por série*

Séries	Alunos
7 <sup>a</sup> série	13,0%
8 <sup>a</sup> série	22,1%
1 <sup>o</sup> ano	45,4%
2 <sup>o</sup> ano	18,4%
3 <sup>o</sup> ano	01,2%

Fonte: Relatório do Pisa, 2003

Os alunos das escolas rurais da Região Norte e os alunos de escolas rurais com menos de 5 alunos não participaram das provas do Pisa.

No terceiro ciclo do Pisa, em 2006, participaram mais de 400.000 estudantes em 57

países, sendo 30 membros da OCDE. Neste ciclo, o foco principal foi em Ciências, mas a avaliação também incluiu Leitura e Matemática. Foram coletados dados sobre os estudantes e suas famílias, e sobre fatores institucionais que podem explicar as diferenças em desempenho.

No Brasil, as provas foram aplicadas em agosto de 2006 em 625 escolas de 390 municípios de todas as unidades da federação. O Brasil apresentou desempenho similar aos anos anteriores em Ciências, uma ligeira queda em Leitura e um aumento significativo em Matemática, que valeu destaque por parte da OCDE. Nesta ciclo do Pisa, o Brasil subiu 14 pontos em Matemática, só superado entre os convidados pela Indonésia (31 pontos) e, considerando os membros da OCDE, pelo México (20 pontos).

Os resultados do Pisa são apresentados através de várias publicações que são encontradas no website <http://www.pisa.oecd.org/>. Essas informações e análises fornecem o perfil do desempenho dos estudantes e indica sua evolução no tempo.

A Tabela 2.2 mostra a média dos alunos brasileiros nos três ciclos do Pisa ocorridos até a data deste trabalho.

Tabela 2.2 *Participação e Proficiências Médias - Brasil*

	<i>Pisa</i> 2000	<i>Pisa</i> 2003	<i>Pisa</i> 2006
No. de Participantes	4.893	4.452	9.295
Ciências	375	390	390
Leitura	396	403	393
Matemática	334	356	370

Fonte: INEP

Esses resultados podem ser vistos na Figura 2.2

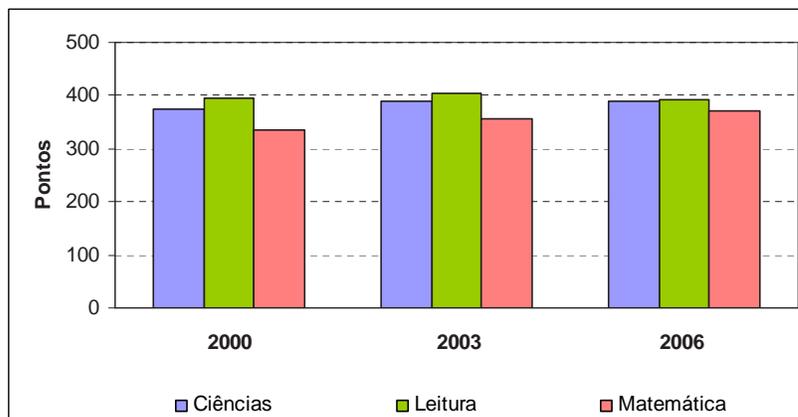


Figura 2.2 Resultado do Pisa no Brasil, nos anos de 2000, 2003 e 2006. Fonte: INEP

## 2.3 Exame Nacional do Ensino Médio (Enem)

Com o propósito de adquirir um mecanismo que pudesse avaliar tanto o aluno quanto a qualidade de ensino de sua escola, o Ministério de Educação (MEC) criou, em 1998, o Exame Nacional do Ensino Médio (Enem). Este exame destina-se aos alunos que já concluíram (egressos) o ensino médio ou vão concluí-lo ao final do ano de realização do exame (concluintes). O Exame é constituído por uma prova única contendo 63 questões objetivas de múltipla escolha (20% de baixo nível de dificuldade, 40% de nível médio e 40% de alto nível de dificuldade) e uma proposta para redação baseada em diferentes textos e linguagens que tratam de temas atuais e nacionais.

As 63 questões da prova objetiva são de igual valor e o total de pontos obtidos é colocado em uma escala de 0 a 100. O instrumento permite, também, que o desempenho em cada uma das cinco competências seja também representado numa escala de 0 a 100.

O objetivo da prova é avaliar as competências e habilidades desenvolvidas pelos participantes ao longo da escolaridade básica, a partir de uma Matriz de Competências desenvolvida para estruturar o exame.

A prova objetiva avalia 5 competências e 21 habilidades (veja site: [www.enem.inep.gov.br](http://www.enem.inep.gov.br)).

Diagrama das cinco competências e 21 habilidades:

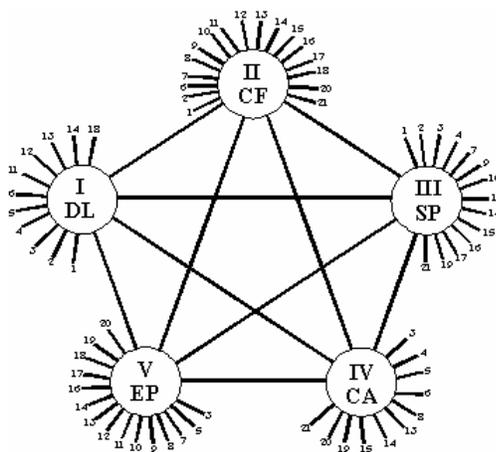


Figura 2.3 Associação entre Competências e Habilidades da Matriz do Enem

- I - Dominar linguagens (DL);
- II - Compreender fenômenos (CF);
- III - Enfrentar situações-problema (SP);
- IV - Construir argumentações (CA); e
- V - Elaborar propostas (EP).

A redação avalia as seguintes competências: 1) Demonstrar domínio da norma culta da língua escrita, 2) Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo, 3) Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista, 4) Demonstrar conhecimento dos mecanismos lingüísticos necessários para a construção da argumentação e 5) Elaborar proposta de intervenção para o problema abordado demonstrando respeito aos direitos humanos.

A maior parte dos participantes do Enem é de egressos. Em 2007 mais de 65% dos

inscritos concluíram o Ensino Médio em anos anteriores a 2007. Em 2007 compareceram ao exame cerca de 2.738.610 candidatos. O desempenho médio neste ano foi 51,52 pontos na parte objetiva da prova, e 55,99 pontos na redação.

As notas do Enem são utilizadas como critério para a distribuição de bolsas do Programa Universidade para Todos (ProUni). Os resultados do Enem servem para diagnosticar a situação do ensino médio do País permitindo que se compare o desempenho dos estudantes concluintes do ensino médio entre diferentes escolas, no mesmo ano (relatório do INEP). Várias instituições de nível superior utilizam o resultado do Enem para o seu processo de seleção. Por questões técnicas, os resultados do Enem não são entre anos.

## **2.4 Exame Nacional de Certificação de Competências para o Ensino de Jovens e Adultos (Encceja)**

O Encceja é o instrumento que avalia as competências e habilidades de jovens e adultos, residentes no Brasil e no exterior, em nível de conclusão do Ensino Fundamental e do Ensino Médio. Foi criado inicialmente para controlar a certificação de brasileiros que cursavam o ensino básico fora do país, especialmente no Japão, onde foi aplicado pela primeira vez em 2002. Na Suíça este exame foi aplicado pela primeira vez em 2005. A adesão ao Encceja é de caráter opcional, sendo também aplicado para brasileiros detidos em instituições do Japão e da Europa.

Os principais objetivos do Encceja são:

1. construir uma referência nacional de auto-avaliação para jovens e adultos por meio de avaliação de competências e habilidades, adquiridas no processo escolar ou nos processos formativos que se desenvolvem na vida familiar, na convivência humana, no trabalho, nos movimentos sociais e organizações da sociedade civil e nas manifestações culturais;
2. estruturar uma avaliação direcionada a jovens e adultos que sirva às Secretarias da Educação para que procedam à aferição de conhecimentos e habilidades dos participantes no nível de conclusão do ensino fundamental e do ensino médio nos termos do artigo 38, §§ 1º e 2º da Lei nº 9.394/96 - Lei das Diretrizes e Bases da Educação Nacional.

Em 2002 participaram do Encceja 14.488 alunos e em 2005, 38.687. O Encceja é elaborado levando-se em consideração as especificidades do aluno (jovens e adultos). A prova do Encceja é uma prova clássica, tem uma parte como o Enem, de conhecimentos gerais, mas também trabalha com competências e habilidades.

#### **2.4.1 Estrutura do Encceja**

As provas do Ensino Fundamental constam das quatro áreas do conhecimento a saber:

1. Língua Portuguesa, Língua Estrangeira, Educação Artística e Educação Física;
2. História e Geografia;
3. Matemática;
4. Ciências Naturais, estabelecida na Base Nacional Comum.

As provas do Ensino Médio constam das quatro áreas do conhecimento a saber:

1. Línguas, Códigos e suas Tecnologias;
2. Ciências Humanas e suas Tecnologias;
3. Matemática e suas Tecnologias e
4. Ciências Naturais e suas Tecnologias, estabelecida na Base Nacional Comum.

#### **2.4.2 Matriz de Competência e Habilidades do Encceja**

A matriz de competências e habilidades que estrutura o Encceja, considera simultaneamente as competências relativas às áreas de conhecimento e as que expressam as possibilidades cognitivas de jovens e adultos para a compreensão e realização de tarefas relacionadas com essas áreas: competências do sujeito. Estas competências referem-se ao domínio de linguagens, compreensão de fenômenos, enfrentamento e resolução de situações-problema, capacidade de argumentação e elaboração de propostas. Estas cinco competências do sujeito são eixos cognitivos, associados às competências apresentadas nas disciplinas e áreas do conhecimento do ensino fundamental e do ensino médio. Destas interações resultam, em cada área, habilidades que serão avaliadas por meio de questões objetivas (múltipla escolha) e pela produção de um texto (redação).

O participante do Enccja pode obter certificação em cada prova separadamente. Em 2006, no Ensino Fundamental, foram aplicados 60.941 avaliações e no no Ensino Médio 105.674. O Enccja 2006 foi estruturado em metodologia da Teoria de Resposta ao Item, sendo que a escala de proficiência adotada variou de 60 a 180, onde o nível 100 por questões técnica, foi sugerido como ponto indicativo de que o participante desenvolveu as habilidades mínimas necessárias para obter a certificação, lembrando que as redes de ensino são livres para definir seu próprio parâmetro.

## **2.5 Sistema de Avaliação da Educação Básica (Saeb)**

O Sistema Nacional de Avaliação da Educação Básica é o mais amplo instrumento de avaliação externa da qualidade do desenvolvimento de habilidades e competências dos estudantes brasileiros. É também um dos mais amplos sistemas de avaliação em larga escala da América Latina. De acordo com a portaria nº 931, de 21 de março de 2005, o Saeb passou a ser composto por duas avaliações: a Avaliação Nacional da Educação Básica (Aneb), que tem foco nas gestões dos sistemas educacionais, e a Avaliação Nacional de Rendimento Escolar (Anresc) que tem foco em cada unidade escolar. A Aneb recebe o nome de Saeb em suas divulgações, pelo fato de manter as características do mesmo, e a Anresc passou a ser chamada de Prova Brasil em suas divulgações. No próximo Capítulo é feito um breve estudo sobre o Saeb.

---

## Capítulo 3

# SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA (Saeb)

---

### 3.1 Introdução

Desde sua primeira edição em 1990, o sistema passou por várias modificações, até sua maior modificação em 2005 que o desdobrou em dois exames. O Aneb é uma avaliação amostral em larga escala, que tem por objetivo contribuir para melhorar a qualidade das escolas, alterando a prática pedagógica e as características gerais do Sistema Nacional. Seu interesse não é avaliar estudantes e nem escolas individualmente, mas o Sistema Nacional como um todo.

O Saeb coleta dados sobre o desempenho dos alunos e as condições intra e extra escolares que neles interferem. A análise dos resultados dos levantamentos do Saeb permite acompanhar a evolução do desempenho dos alunos, bem como dos diversos fatores que interferem na qualidade do ensino ministrado nas escolas. Esses resultados são utilizados pelos professores, administradores e gestores da educação, para que possam avaliar se os alunos estão realmente adquirindo as habilidades e os conhecimentos necessários à sua plena inserção na sociedade.

Neste Capítulo será feito um pequeno relato histórico do Saeb, bem como uma breve descrição da Prova Brasil e do Ideb (Índice de Desenvolvimento da Educação Básica).

### 3.2 Histórico

O Saeb foi criado pelo Instituto Nacional de Estudos e Pesquisas Educacionais (INEP) tendo seu primeiro levantamento em 1990, quando foram avaliados somente alunos do Ensino Fundamental pertencentes à rede pública.

Em 1993, foi realizada a segunda aplicação do Saeb, estruturada em três eixos de estudo:

(1) rendimento do aluno; (2) perfil e prática docentes; (3) perfil dos diretores e formas de gestão escolar.

As amostras utilizadas pelo Saeb nesses dois primeiros ciclos foi formada por alunos das 1<sup>a</sup>, 3<sup>a</sup> e 7<sup>a</sup> séries de escolas da rede pública de ensino. A partir do 3<sup>o</sup> ciclo de aplicação, em 1995, o Saeb fez as seguintes modificações no seu plano de trabalho: incluiu em sua amostra o ensino médio e a rede particular de ensino; adotou técnicas mais modernas de medição do desempenho dos alunos; incorporou instrumentos de levantamento de dados sobre as características socioeconômicas e culturais e sobre os hábitos de estudo dos alunos e redefiniu as séries avaliadas, selecionando aquelas conclusivas de um determinado ciclo escolar: 4<sup>a</sup> e 8<sup>a</sup> séries do ensino fundamental e 3<sup>a</sup> série do ensino médio, e no ciclo de 1997 introduziu a construção das Matrizes de Referência. No ciclo de 1999 foi adotado o mesmo modelo de 1997. No 6<sup>o</sup> ciclo do Saeb, em 2001 foi feita atualização das Matrizes de Referência.

O 7<sup>o</sup> ciclo, aplicado em 2003, trouxe uma inovação que foi a de acompanhar o desempenho dos alunos que participam do programa Bolsa-Escola do Ministério da Educação, outra inovação foi sobre o questionário socioeconômico, aplicado a alunos, professores e diretores, que incluiu questões sobre o problema da violência. Um grupo de estudo aprofundou uma análise sobre a cor e o rendimento dos estudantes. Nesse ano, o universo investigado foi formado por alunos pertencentes a escolas listadas no Censo Escolar de 2002 que estavam freqüentando as séries a serem avaliadas neste ano. Foram feitas as seguintes modificações, de acordo com o Inep, para 2003:

- a) todos os alunos matriculados em 2003 nas escolas urbanas constantes no Censo Escolar de 2002 em uma das três séries de interesse (a 4<sup>a</sup> e a 8<sup>a</sup> séries do ensino fundamental e a 3<sup>a</sup> série do ensino médio) exceto os das turmas multisseriadas, de aceleração e da educação de jovens e adultos.
- b) foram incluídos os alunos da 4<sup>a</sup> série das escolas rurais com 10 ou mais alunos nesta série e ainda os alunos das escolas federais urbanas das três séries avaliadas.
- c) as escolas urbanas foram separadas em grupos (estratificadas) segundo três conjuntos de municípios:

- os da região metropolitana sediada na capital do estado (sempre que houver) ou para o município da capital se não houver região metropolitana;
  - para o conjunto dos municípios não metropolitanos com população superior a 200.000 habitantes (se houver) e
  - para o restante dos municípios.
- d) estes três conjuntos de municípios foram denominados de Região Metropolitana ou Município da Capital, Grandes Cidades Não Metropolitanas e Outras Cidades.
- e) os alunos da 4ª série das escolas rurais serão mantidos em um grupo de análise à parte.

Em 2003, participaram do Saeb cerca de 300 mil alunos, 17 mil professores e 6 mil diretores de 6.270 escolas das 27 unidades da Federação. Além desses, perto de 4 mil pessoas, entre aplicadores dos testes, supervisores e funcionários das Secretarias de Educação Estaduais trabalharam nessa aplicação.

Em 2005 foi aplicado o 8º ciclo do Saeb e a novidade foi a criação da Aneb e da Anresc, conforme foi visto no Capítulo anterior. Em novembro de 2007 aconteceu mais um levantamento do Saeb.

O Saeb avalia estudantes de escolas urbanas e rurais, tanto da rede pública quanto da rede privada. A participação é amostral, portanto, com resultados disponíveis em esfera nacional, regional e por unidade da Federação, para as séries e disciplinas avaliadas, sem detalhamento para municípios ou unidades de ensino.

O Saeb é aplicado a cada dois anos em amostras de alunos da 4<sup>a</sup> e 8<sup>a</sup> séries do Ensino Fundamental e 3<sup>a</sup> série do Ensino Médio, representativas do País e de todas as Unidades da Federação. Os alunos fazem provas de Língua Portuguesa (Foco: Leitura) e Matemática (Foco: resolução de problemas). Eles também respondem a um questionário sobre seus hábitos de estudo e suas características socioculturais. Os professores e diretores participam, respondendo a questionários que informam sobre perfil e prática docente, mecanismos de gestão e infra-estrutura da escola.

O Saeb adota como referência a média nacional dos alunos da 8<sup>a</sup>série de 1997, que é de 250 pontos. A partir das premissas da Teoria da Resposta ao Item (TRI) é possível construir uma escala única de pontuações, desenvolvendo matrizes de referência para populações de distintos níveis de escolaridade: no caso do Brasil, 4<sup>a</sup> e 8<sup>a</sup> séries do Ensino Fundamental e 3<sup>a</sup> do Ensino Médio. Isto permite comparar as médias de “proficiência” em cada disciplina entre os diversos níveis do sistema educativo e entre as regiões do país, e entre os vários anos, colocando todos os níveis em uma mesma escala. A partir das informações do Saeb, o MEC e as Secretarias Estaduais e Municipais de Educação podem definir ações voltadas para a correção de distorções e debilidades identificadas e direcionar seus recursos técnicos e financeiros para áreas prioritárias, visando ao desenvolvimento do Sistema Educacional Brasileiro e à redução das desigualdades. Algumas das metodologias utilizadas pelo Saeb serão vistas a seguir.

### 3.3 Metodologias do Saeb

Dentre as várias metodologias utilizadas pelo Saeb destacam-se: Adoção de Matrizes de Referência, Blocos Incompletos Balanceados (BIB), Teoria Clássica dos Testes, Análise Fatorial, Amostragem Complexa, Modelo Linear Hierárquico, Teoria de Resposta ao Item (TRI) e Escalas de Proficiência utilizadas para interpretar e descrever o desempenho dos alunos.

#### 3.3.1 Matrizes de Referência

As Matrizes de Referência são um documento no qual estão descritas as orientações para a elaboração dos itens (questões que compõem a prova) dos testes do Saeb sendo portanto

a base para a elaboração destes testes. Elas têm por referência os Parâmetros Curriculares Nacionais.

Para compor as Matrizes de Referência foram realizada consultas às propostas curriculares dos Estados brasileiros, obtendo-se uma síntese do que havia de comum entre elas. Para validar esta síntese consultaram-se os professores das capitais brasileiras regentes das redes municipal, estadual e privada na 4ª e 8ª série do Ensino Fundamental e na 3ª série do Ensino Médio, nas disciplinas de Língua Portuguesa e Matemática e examinaram-se os livros didáticos mais utilizados nas mesmas redes e séries. Em seguida, foram incorporadas análises de professores e especialistas nas áreas do conhecimento avaliadas pelo Saeb. Decorrente destas análises, a opção teórica adotada é a que pressupõe a existência de competências cognitivas e habilidades a serem desenvolvidas pelo aluno no processo de ensino-aprendizagem.

Cada Matriz de Referência é estruturada em tópicos e respectivos descritores que indicam as competências e habilidades de Língua Portuguesa e Matemática a serem avaliadas.

As Matrizes de Referência para Matemática são limitadas, pois algumas habilidades e competência não podem ser medidas por meio de um teste escrito, como por exemplo: *Utilizar Procedimentos de Cálculo Mental* e *Construir Representações Gráficas tais como Listas, Tabelas e Gráficos*, que apesar de serem indicadores importantes, não possuem descritores correspondentes nessa Matriz. No Apêndice A estão as Matrizes de Referência para Língua Portuguesa e Matemática, usadas pelo Saeb.

Por meio dos itens do teste do Saeb pode-se dizer que um aluno desenvolveu uma determinada habilidade (constante em um descritor) quando ele é capaz de resolver um problema a partir da utilização e aplicação de um conceito por ele já construído. Procura-se apresentar, por exemplo, ao aluno, situações em que a resolução de problemas seja significativa para ele.

### 3.3.2 Testes Padronizados

O Saeb aplica testes padronizados para descrever o que os alunos sabem e são capazes de fazer nas disciplinas de Língua Portuguesa e Matemática, em momentos conclusivos de seu percurso escolar (4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio).

Esses testes são de múltipla escolha, elaborados a partir de descritores das Matrizes de Referência para o Saeb.

Durante a realização dos testes do Saeb é aplicado um questionário que coleta informações sobre o contexto social, econômico e cultural dos alunos, e ainda sobre a trajetória de sua escolarização, buscando apresentar indicações do efeito que alguns destes fatores têm sobre o desempenho. Professores e diretores também são convidados a responder questionários que possibilitam conhecer a formação profissional, práticas pedagógicas, nível socioeconômico e cultural, estilos de liderança e formas de gestão.

### 3.3.3 Bloco Incompleto Balanceado (BIB)

Essa metodologia permite a aplicação de 169 itens de forma a cobrir a Matriz de Referência em cada série e disciplina. Esse conjunto de itens é dividido em 13 blocos, com 13 itens cada, agrupando-os de três em três, em 26 cadernos diferentes de prova. Um modo de fazer a distribuição dos cadernos encontra-se na Tabela 3.1.

Tabela 3.1 *Bloco Incompleto Balanceado - BIB*

Caderno	A	B	C	Caderno	A	B	C
<b>1</b>	1	2	5	<b>14</b>	1	3	8
<b>2</b>	2	3	6	<b>15</b>	2	4	9
<b>3</b>	3	4	7	<b>16</b>	3	5	10
<b>4</b>	4	5	8	<b>17</b>	4	6	11
<b>5</b>	5	6	9	<b>18</b>	5	7	12
<b>6</b>	6	7	10	<b>19</b>	6	8	13
<b>7</b>	7	8	11	<b>20</b>	7	9	1
<b>8</b>	8	9	12	<b>21</b>	8	10	2
<b>9</b>	9	10	13	<b>22</b>	9	11	3
<b>10</b>	10	11	1	<b>23</b>	10	12	4
<b>11</b>	11	12	2	<b>24</b>	11	13	5
<b>12</b>	12	13	3	<b>25</b>	12	1	6
<b>13</b>	13	1	4	<b>26</b>	13	2	7

Fonte: Inep - Saeb

onde A, B e C representam os Blocos.

Desse modo, é avaliado um amplo escopo de conteúdos, onde cada aluno responde apenas 39 itens. Para comparar o desempenho dos alunos entre as séries avaliadas aplicam-se

blocos da 4ª série do ensino fundamental na 8ª série do ensino fundamental, bem como blocos da 8ª série do ensino fundamental na 3ª série do ensino médio. Para poder comparar as séries históricas, mantêm-se alguns blocos comuns e/ou itens já aplicados em anos anteriores. A análise dos resultados é feita utilizando-se a Teoria da Resposta ao Item, que permite a comparação e a colocação dos mesmos em uma escala única de desempenho dos alunos nas áreas selecionadas, ainda que estes tenham respondido a diferentes conjuntos de itens.

### 3.3.4 Escalas de Proficiência

A Escala de Proficiência ou de desempenho, que descreve em cada nível as competências e as habilidades que os alunos são capazes de demonstrar, é única para cada disciplina e permite apresentar, em uma mesma métrica, os resultados de desempenhos dos estudantes de todas as séries (4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio) e anos de aplicação dos testes. Um estudo sobre as “Escalas” utilizadas pelo Saeb é feito no Capítulo 6.

### 3.3.5 Teoria da Resposta ao Item

Teoria de Resposta ao Item (TRI), ferramenta base deste trabalho, é uma metodologia baseada em um conjunto de modelos matemáticos que torna possível comparar o desempenho dos alunos em diferentes testes. Ela surge da necessidade de superar as limitações da apresentação dos resultados somente através de percentual de acertos, médias ou escores de testes. A partir da TRI tornou-se possível posicionar os alunos em escalas comuns de proficiências, mesmo que nem todos tenham respondido aos mesmos itens de teste de avaliação.

Mais detalhes sobre essa teoria será visto no Capítulo 5.

## 3.4 Resultados do Saeb

Os resultados do Saeb são apresentados de duas formas: mostrando o comportamento no tempo, como pode ser visto no Apêndice B, podendo-se consultar o sitio do INEP ([www.inep.gov.br](http://www.inep.gov.br)) para obter os demais resultados, ou tomando-se o percentual de estudantes distribuídos em quatro intervalos de proficiência, denominados de *muito crítico*,

*crítico, intermediário e adequado.* Em 2007 essas nomenclaturas poderão ser modificadas, porém até a data desta Dissertação ainda não foram divulgados os resultados do Saeb 2007.

### 3.5 A Prova Brasil

Criado em 2005 por demanda das escolas que desejavam ter seus resultados próprios, a prova Brasil (Anresc) passou a ter dimensões gigantescas, avaliando cerca de 3 milhões de estudantes naquele ano e devendo chegar a 5 milhões em 2007. Os resultados dessa avaliação são usados pelo MEC como política de distribuição de recursos através do Índice de Desenvolvimento da Educação Básica (Ideb). Em termos de estrutura, o exame será composto por 21 tipos de prova (cadernos). Cada caderno é constituído por quatro blocos, sendo que dois são destinados a respostas de Língua Portuguesa e os outros dois abordam questões de Matemática. Os testes são de múltipla escolha, com quatro ou cinco alternativas de resposta para cada questão, sendo que apenas uma está correta. Os alunos de 4ª série responderão a 22 itens de português e a 22 itens de matemática. Já os estudantes de 8ª série e do 3º ano do ensino médio responderão a 26 itens de português e a 26 de matemática. Toda a metodologia de construção dos resultados, através da Teoria da Resposta ao Item, é equivalente ao do Saeb (Aneb).

### 3.6 O Ideb

Comparativamente aos resultados do Pisa, nota-se que o Brasil está muito aquém do desejado, cerca de um desvio-padrão, em termos de proficiência. Por conta disso foi criado um índice no intervalo [0,10] visando dar à escola uma ferramenta de melhor acompanhamento.

O Ideb é um indicador de qualidade educacional que combina informações de desempenho em exames padronizados (Prova Brasil ou Saeb) - obtido pelos estudantes ao final das etapas de ensino (4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio) - com informações sobre rendimento escolar (aprovação). Para maiores detalhes sobre a definição e construção do Ideb, consultar a publicação Índice de Desenvolvimento da Educação Básica (Ideb), na Série Documental - Texto para Discussão nº 26, disponível em: [www.inep.gov.br](http://www.inep.gov.br).

Antes da análise dos itens pela Teoria da Resposta ao Item, o Saeb faz uma análise

clássica dos mesmos. No Capítulo 4 apresenta-se um referencial teórico sobre a Teoria Clássica dos Testes.

## TEORIA CLÁSSICA DOS TESTES

---

### 4.1 Introdução

A Psicometria constitui uma das várias formas de medição em Psicologia (Pasquali, 2003, p.53), assumindo símbolos que expressam parâmetros, os quais representam variáveis de caráter abstrato, suficientes para o desenvolvimento do modelo matemático para essa teoria. Em ciências psicossociais destacam-se duas formas de medida: medida por lei e medida por teoria. No caso de medida por teoria destaca-se *Teoria Psicométrica ou Teoria dos Testes Psicológicos*. Os parâmetros com os quais a *Teoria dos Testes Psicológicos* trabalha são: a resposta (comportamento) do sujeito e o critério. Ela procura explicar o sentido que têm as respostas dadas pelos sujeitos a uma série de tarefas. O critério sendo entendido de várias maneiras possibilitou o nascimento de duas teorias psicométricas: a Teoria Clássica dos Testes (TCT), que entende o critério como comportamento (futuro) e a Teoria de Resposta ao Item (TRI), que entende como critério o traço latente (construto, habilidade, proficiência). A primeira define a qualidade dos testes psicológicos (variáveis observáveis) em termos de um critério representado por *comportamentos* (presentes ou futuros) do ser humano. A segunda define a qualidade dos testes psicológicos (variáveis observáveis) em função de um critério que são variáveis hipotéticas, geralmente denotadas por  $\theta$ , ou traço latente (ver Pasquali, 2003, p.55). O interesse da TCT está centrado no escore total do examinando no teste (soma dos escores nos itens). Os itens elaborados para cada prova podem ser analisados qualitativamente e quantitativamente. A análise quantitativa inclui procedimentos da Teoria Clássica dos Testes, da Análise Fatorial e da Teoria da Resposta ao Item. Neste Capítulo será feita uma breve revisão teórica da Teoria Clássica dos Testes.

## 4.2 O Modelo da Psicometria Clássica: TCT

O modelo clássico da psicometria tradicional (Pasquali, 2003) está fundamentado na Teoria Clássica dos Testes. Esta teoria está apoiada no seguinte paradigma: o escore empírico ou bruto do sujeito é constituído de dois componentes: 1) o escore real ou verdadeiro ( $V$ ) do sujeito no comportamento avaliado; e 2) o erro de medida ( $E$ ). O erro, sempre presente em qualquer medida empírica, resulta no modelo fundamental da psicometria, o qual confirma a tese de que o escore bruto de um examinando é a soma do escore verdadeiro e do erro ( $T = V + E$ ). Este modelo implica alguns postulados básicos: a) o escore esperado é o escore verdadeiro. Isto decorre do conceito de esperança matemática do escore empírico, ou seja, se o sujeito responde infinitas vezes ao mesmo teste, ele terá infinitos diferentes escores empíricos, e a média destes infinitos escores será o escore verdadeiro, porque ela eliminaria os erros; b) não há correlação entre o escore verdadeiro e o erro, ou seja, é zero; portanto, não há nenhuma razão para supor que escores verdadeiros maiores terão erros positivos e escores verdadeiros menores terão erros negativos; e c) os erros em testes paralelos não são correlacionados (Pasquali, 2003).

## 4.3 Análise Algébrica dos Itens

Tratando-se da TCT, na análise algébrica do item são avaliados: *a unidimensionalidade*, *a dificuldade*, *a discriminação*, *os vieses*, *a validade* e *a precisão*. Veremos algumas destas. Para Pasquali (2003), neste tipo de análise supõe-se que quando se está analisando uma série de itens, todos eles estão medindo uma mesma coisa (unidimensionalidade). Essa é uma condição necessária para analisar qualquer característica ulterior de um item, como a dificuldade e a discriminação do mesmo. A dificuldade do item é definida em termos de percentagem de sujeitos que dão a resposta certa a um item. A discriminação do item é a sua capacidade de diferenciar sujeitos com escores altos no teste de sujeitos com escores baixos. Para calcular discriminação pode-se usar dois procedimentos: a) a dos grupos-critério e b) a correlação do item com o total dos itens. Existem dois tipos de critérios: os internos e os externos. Utilizam-se critérios internos ao próprio grupo para definir estes grupos-critério (ver Pasquali, 2003). Existem diversas técnicas estatísticas para determinar o índice de discriminação do item (ver Anastasi, 1988), como por exemplo a correlação item total, a correlação ponto-bisserial, a correlação bisserial e a correlação tetracórica.

A primeira análise feita pelo Saeb é a análise clássica dos itens. São calculados os percentuais de acerto do item e o percentual de escolha em cada alternativa, o índice de discriminação, o percentual de acerto do grupo superior e inferior pelos 27% de alunos com maiores e menores medidas de desempenho, respectivamente, e o coeficiente bisserial do item e por alternativa.

O coeficiente bisserial, representado por  $r_{bis}$ , é uma medida de associação entre o desempenho no item e o no teste, que estima a correlação entre a variável de desempenho no teste e uma variável latente com distribuição normal, que por hipótese, representa a *habilidade* que determina o acerto ou erro no item. A fórmula que expressa o coeficiente bisserial utilizada pelo Saeb (Relatório de Análise Clássica do Teste) é:

$$r_{bis} = \frac{M^+ - M}{S} \times \frac{p}{h(p)} = \frac{M^+ - M^-}{S} \times \frac{p(1-p)}{h(p)}, \quad (4.1)$$

onde

$M^+$  é a média da medida de desempenho no teste para os alunos que acertaram o item;

$M$  é a média geral da medida de desempenho no teste para todos os alunos;

$M^-$  é a média da medida de desempenho no teste para os alunos que erraram o item;

$S$  é o desvio padrão da medida de desempenho no teste para todos os alunos;

$p$  é o percentual de acerto no item;

$h(p)$  é o valor da densidade da distribuição normal com média 0 (zero) e variância 1 (um) no ponto em que a área da curva à esquerda deste ponto é igual a  $p$ .

As correlações bisserias feitas pelo Saeb foram calculadas com o *normit*, uma variável obtida por uma transformação não linear a partir dos escores dos examinados. O escore *normit* de um aluno que acerta  $i$  itens de um caderno é definido por:

$$normit(i) = \phi^{-1} \left( \frac{CF(i) + CF(i-1)}{2n} \right), i > 0 \quad (4.2)$$

onde:

$\phi$  é a função cumulativa da distribuição normal com média 0 e variância 1;

$CF(i)$  é a frequência cumulativa para  $i$  itens corretos em um caderno e

$N$  é o número de alunos que responderam o caderno. Se não há nenhum acerto, então  $i=0$ , logo o *normit(0)* será:

$$\text{normit}(0) = \phi^{-1} \left( \frac{CF(0) + \frac{CF(1)}{2}}{2N} \right)$$

Na Teoria Clássica dos Testes surgem problemas do tipo: “os parâmetro clássicos (dificuldade e discriminação) dependem diretamente da amostra de sujeitos utilizada para estabelecê-los, e se a amostra não for rigorosamente representativa da população, aqueles parâmetros dos itens não podem ser considerados válidos para esta população”. Outro problema da Teoria Clássica dos Testes consiste em que ela é orientada para o teste total e não para o item individual. Para sanar estes e outros problemas surge uma teoria chamada Teoria de Resposta ao Item, que está apresentada no capítulo seguinte.

## TEORIA DA RESPOSTA AO ITEM

---

### 5.1 Introdução

Imagine que numa prova com 30 itens (questões) de múltipla escolha, o aluno X acertou 15 questões mais fáceis e o aluno Y acertou 15 questões mais difíceis. Pela TCT esses alunos terão o mesmo escore, pois essa Teoria considera a prova como um todo. Mas será que atribuir o mesmo escore para esses dois alunos é justo? Já a Teoria da Resposta ao Item (TRI) analisa cada item individualmente, sendo possível saber qual dos dois alunos tem maior proficiência. Ela trabalha com os traços latentes (variáveis não observáveis) dos respondentes e tem como elementos centrais da análise os *itens*; desse modo, a análise de um teste feita pela TRI tem melhor validação do que a análise feita pela TCT.

Uma vantagem da TRI em relação a TCT pode ser verificada através do seguinte exemplo: Deseja-se comparar o nível de conhecimento entre alunos de duas séries distintas (5<sup>a</sup> e 8<sup>a</sup> séries do ensino fundamental, por exemplo). Na TCT esta comparação somente será possível caso seja aplicada a mesma prova para as duas turmas, enquanto que na TRI são necessárias apenas algumas questões em comum nas provas aplicadas às duas séries.

A TRI é um instrumento que, atualmente, vem sendo aplicado nos processos quantitativos na área da avaliação educacional, tendo sido aplicada pela primeira vez no Brasil em 1995, na análise dos resultados do Sistema Nacional de Ensino Básico (Saeb). Além da aplicação da TRI na área educacional, ela também pode ser aplicada em outras áreas como: psicossocial, médica e marketing. A TRI baseia-se em um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente (Andrade, Tavares & Valle, 2000), ou seja, o que a TRI propõe são modelos de variáveis latentes para representar a relação entre a probabilidade de um respondente apresentar determinada resposta a um item e seus traços latentes ou habilidades na área do

conhecimento avaliada, os quais não podem ser observados diretamente. Essas habilidades freqüentemente são tratadas por proficiências ou aptidões.

O termo *grupo*, utilizado neste trabalho, corresponde a uma amostra de indivíduos de uma *população*. Um único grupo de respondentes significa que a amostra de indivíduos foi retirada de uma mesma população, enquanto que dois grupos (ou mais) de respondentes são dois conjuntos distintos de indivíduos, que foram amostrados de duas (ou mais) populações. Na TRI trabalha-se com modelos para itens dicotômicos ou dicotimizados e com modelos para itens não dicotômicos como por exemplo: Modelo de Resposta Nominal, Modelo de Resposta Gradual, Modelo de Crédito Parcial, Modelo de Crédito Parcial Generalizado. Nesta Dissertação apresenta-se apenas os modelos unidimensionais mais utilizados para um único grupo de respondentes, ou sejam os que consideram *somente um traço latente (ou uma habilidade)*, e estuda-se como fazer a estimação de parâmetros.

## 5.2 Modelos Matemáticos da Teoria Resposta ao Item

Os modelos propostos pela TRI dependem basicamente de três fatores, a saber: da natureza do item (que pode ser dicotômico ou não dicotômico), do número de populações envolvidas (que pode ser apenas uma ou mais de uma população) e da quantidade de traços latentes que está sendo medida (que pode ser apenas um ou mais de um traço latente), (ver Andrade, Tavares e Valle, 2000). Modelos que avaliam apenas um traço latente são chamados de modelos unidimensionais e, obviamente, quando avaliam mais de um traço latente são chamados de modelos multidimensionais.

Pode-se classificar os modelos envolvendo um único grupo em: *modelos dicotômicos ou dicotimizados* e os modelos para *itens não dicotômicos*.

Os modelos envolvendo duas ou mais populações são basicamente extensões dos modelos envolvendo um único grupo, sendo que poucos modelos já se encontram implementados computacionalmente.

Na área de Avaliação Educacional pode-se definir uma população por uma determinada característica que depende do objetivo do estudo. Se, por exemplo, o estudo for sobre a 5.<sup>a</sup> série do Ensino Fundamental de um determinado lugar, então toma-se uma única

amostra dos alunos dessa população, composta de alunos do período diurno e do noturno. Note que neste caso existe apenas um único grupo de respondentes. Mas, caso o interesse fosse sobre a 5.<sup>a</sup> série *diurna* e a 5.<sup>a</sup> série *noturna* do Ensino Fundamental desse mesmo lugar ou de qualquer outro, seriam tomadas duas amostras: uma dos alunos do período diurno e outra dos alunos do noturno. Teria-se então dois grupos de alunos. Portanto, é pelo próprio processo de amostragem do estudo que define quantas (e quais) populações estão envolvidas. Apresenta-se a seguir alguns modelos que são utilizados quando um teste é aplicado a um único grupo de respondentes (adaptado de Andrade, Tavares e & Valle, 2000). Será feita a revisão teórica dos seguintes modelos matemáticos para itens dicotômicos ou dicotomizados: Modelo de Rasch, Modelo Logístico de Um Parâmetro (ML1), Modelo Logístico de Dois Parâmetros (ML2) e Modelo Logístico de Três Parâmetro (ML3) e também das Curvas características do Item e da Função de Informação do Item. Será feita também a revisão teórica do processo de estimação conhecido por *Máxima Verossimilhança Marginal*.

A TRI utiliza duas funções matemáticas para caracterizar os parâmetros dos itens componentes de um teste: a função distribuição da normal padronizada (ou função Gaussiana ou ogiva normal) e a função logística. Nestas funções a proficiência do indivíduo é normalmente representada por  $\theta$ , e os parâmetros do item  $i$  por  $a_i$ ,  $b_i$  e  $c_i$ .

Os primeiros modelos de resposta ao item surgiram na década de 50. Nestes modelos considerava-se que uma única habilidade de um único grupo, estava sendo medida por um teste onde os itens eram corrigidos de maneira dicotômica (Andrade, Tavares & Valle, 2000). Esses modelos foram primeiramente desenvolvidos na forma de uma função distribuição da normal, ou seja:

$$\Phi(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt, \quad (5.1)$$

e depois foram descritos por uma forma mais conveniente computacionalmente que é a função logística (que é uma função explícita dos parâmetros dos itens):

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1 + e^{D(\theta-b_i)}}, \quad (5.2)$$

Ressalta-se que a função (5.1) envolve uma integral que precisa ser resolvida numerica-

mente no processo de estimação dos parâmetros dos itens e habilidades, e diversas vezes, incluindo uma fonte adicional de imprecisão no processo.

Lord, em 1952, foi o primeiro a desenvolver o modelo unidimensional de dois parâmetros, baseado na distribuição normal acumulada, citada em (5.1). Após algumas aplicações práticas desse modelo, sentiu a necessidade de introduzir um parâmetro que tratasse do acerto casual (chamado de chute), surgindo assim o modelo de 3 parâmetros.

Em 1968, Birnbaum substituiu em ambos os modelos, a função ogiva normal pela função logística (5.2), função explícita dos parâmetros dos itens e da habilidade dos respondentes e não envolve integração, sendo portanto muito mais fácil de ser trabalhada. A Figura 5.1 mostra que fazendo  $D=1,7$  na função logística as curvas referentes às funções normal e logística praticamente coincidem, e a Figura 5.2 mostra a diferença entre essas duas funções, onde notamos que a diferença absoluta máxima é menor que 0,01.

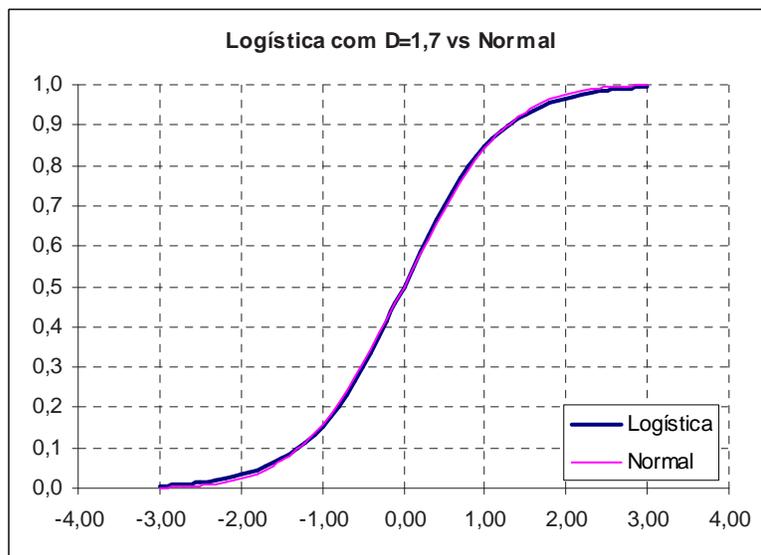
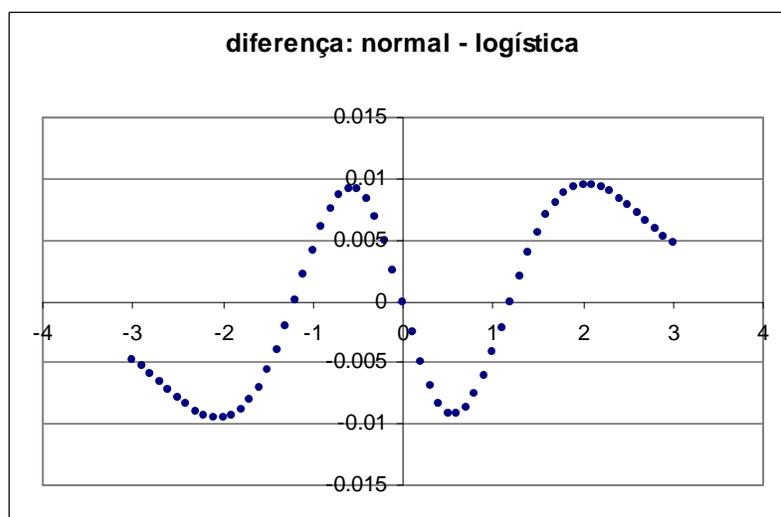


Figura 5.1 *Logística com  $D=1,7$  vs Normal*

Figura 5.2 *Diferença: Normal-logística*

Em 1997, Book & Zimowski introduziram os modelos logísticos de 1, 2 e 3 parâmetros para duas ou mais populações de respondentes.

Os modelos matemáticos apresentados a seguir podem ser utilizados tanto para a análise de itens de múltipla escolha dicotomizados (corrigidos como certo ou errado) quanto para a análise de itens abertos (de resposta livre), quando avaliados de forma dicotomizada. Os modelos da TRI baseam-se em dois postulados: a “Unidimensionalidade” e a “Independência Local”, intimamente relacionados:

*-Unidimensionalidade*

Pelo postulado da Unidimensionalidade *apenas uma aptidão é responsável por um conjunto de tarefas*, mas como para realizar qualquer tarefa o ser humano utiliza-se de várias aptidões (habilidades), para satisfazer o postulado da Unidimensionalidade, é suficiente considerar que exista uma aptidão dominante. Os modelos multidimensionais (que consideram mais de uma aptidão) fazem parte dos modelos pertencentes a Teoria Multidimensional de Resposta ao Item e não fazem parte dos propósitos dessa dissertação.

*-Independência Local*

Este postulado diz que para uma dada habilidade as respostas aos diferentes itens da prova são independentes (Andrade, Tavares & Valle, 2000). Em outras palavras, a habilidade do respondente é informação suficiente para responder a um item, não dependendo das respostas a outros itens; isso, no fundo, implica que o respondente não aprende durante o teste, ficando constante sua habilidade.

**5.2.1 Modelo Logístico de Um Parâmetro (ML1)**

Costuma-se representar o modelo logístico de um parâmetro pela expressão:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}, \quad (5.3)$$

com  $i=1, 2, \dots, I$  e  $j=1, 2, \dots, n$

onde:

$U_{ij}$  representa uma variável dicotômica que assume os valores 1, quando o indivíduo  $j$  responde corretamente o item  $i$ , ou 0 quando o indivíduo  $j$  não responde corretamente o item  $i$ .

$\theta_j$  representa a habilidade do  $j$ -ésimo indivíduo.

$P(U_{ij} = 1|\theta_j)$  é a probabilidade de um indivíduo  $j$  com habilidade  $\theta_j$  responder corretamente o item  $i$ . Essa probabilidade é chamada de Função de Resposta do Item e sua representação gráfica é conhecida por *Curva Característica do Item* para o ML1 e é representada por *CCI*.

$D$  é um fator de escala.

Na Figura (5.3) temos que  $b_1 < b_2$ . Também notamos que para ter uma probabilidade de 50% de responder corretamente ao item precisa-se de uma maior habilidade ( $\theta = b_2$ ) de forma que o item 2 é mais difícil que o item 1, indicando que quanto maior o valor de  $b$ , mais difícil é o item. O parâmetro  $\theta$  também conhecido por parâmetro da habilidade, é uma variável que não pode ser medida diretamente. Este parâmetro representa o traço latente e teoricamente pode assumir valores de  $-\infty$  à  $+\infty$ . Para a construção de escala deve-se estabelecer uma origem e uma unidade de medida, sendo que é costume representar estes valores, respectivamente pela média ( $\mu$ ) e pelo desvio padrão ( $\sigma$ ). Uma escala usada com frequência na TRI é aquela onde  $\mu=0$  e  $\sigma=1$ .

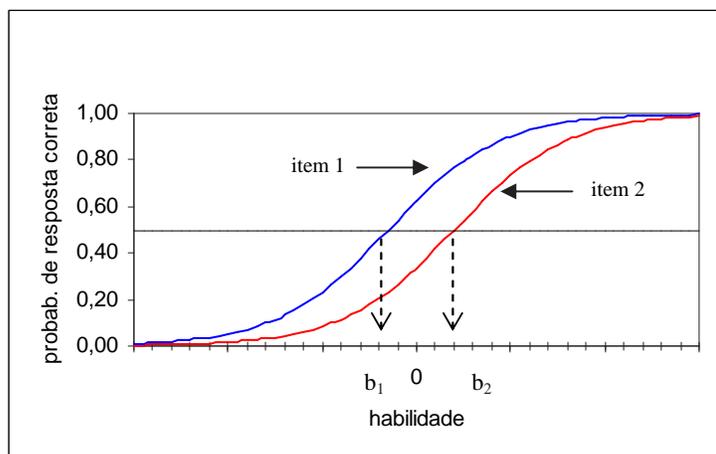


Figura 5.3 *Curva Característica de dois Itens típicos do ML1*

### Modelo de Rasch

Na construção deste modelo, Rasch convencionou adotar “1” para as respostas corretas e “0” para as incorretas. Este modelo expressa a probabilidade de se observar uma resposta correta ao item como função da diferença entre a habilidade ( $\theta_j$ ) do respondente e a dificuldade  $b_i$  do item. Esta relação é representada algebricamente pela fórmula abaixo:

$$P_{ij} = \frac{\exp(\theta_j - b_i)}{1 + \exp(\theta - b)} \quad (5.4)$$

de onde obtém-se

$$P = \exp(\theta - b)(1 - P)$$

e, consecutivamente,

$$\ln\left(\frac{P}{1 - P}\right) = \theta - b, \quad (5.5)$$

Quando a habilidade do respondente for igual a dificuldade do item obtém-se:

$$P_i = \frac{1}{1+1} = \frac{1}{2} = 50\%$$

Se  $\theta > b \Rightarrow P > 0,5$  (probabilidade acima de meio e no máximo igual a 1)

Se  $\theta < b \Rightarrow P < 0,5$  (probabilidade abaixo de meio e no mínimo igual a zero)

### 5.2.2 Modelo Logístico de Dois Parâmetros (ML2)

Este Modelo foi idealizado para avaliar dois parâmetros do item: a dificuldade ( $b_i$ ) e a discriminação ( $a_i$ ), que teoricamente podem assumir um valor real qualquer compreendido entre  $-\infty$  à  $\infty$ . Entretanto para o parâmetro de discriminação considera-se o intervalo de 0 a  $+\infty$ , mas tipicamente o intervalo de 0 a 2. Nesse modelo os valores negativos de  $a_i$  não são considerados, pois indicam que o item é respondido corretamente por respondentes de menor habilidade e errado pelos de maior habilidade o que é um absurdo. Ainda, parâmetro  $a_i$  é proporcional à inclinação no ponto de inflexão,  $b_i$ .

A função logística utilizada para avaliar a probabilidade de um respondente  $j$  com habilidade  $\theta_j$  responder corretamente ao item  $i$  de um teste é representada por:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (5.6)$$

com  $i = 1, 2, \dots, I$ , e  $j = 1, 2, \dots, n$ .

onde  $a_i$  representa o parâmetro de discriminação (ou de inclinação) do item e  $n$  é o número de respondentes. Os demais elementos de (5.6) são os mesmos vistos anteriormente.

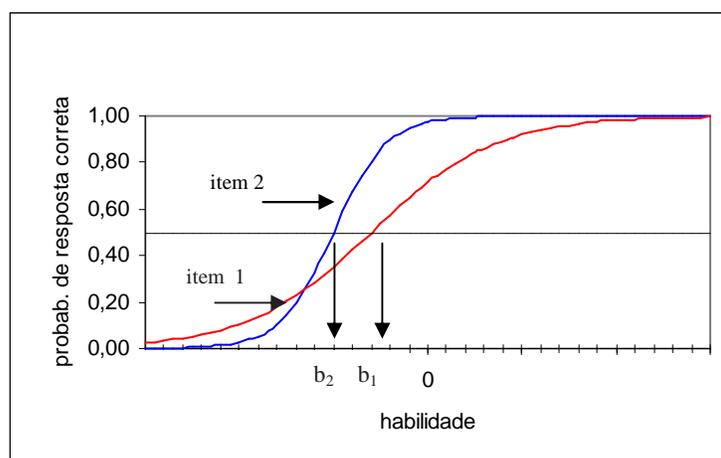


Figura 5.4 Curva Característica de dois Itens típicos do ML2

A Figura (5.4) mostra as CCIs de dois itens que diferem em dois parâmetros: dificuldade (representada por  $b_1$  e  $b_2$ ) e discriminação (que são as inclinações das CCIs dos itens 1 e 2 nos seus respectivos pontos de inflexão).

### 5.2.3 Modelo Logístico de Três Parâmetros (ML3)

O Modelo Logístico de Três Parâmetros é o modelo mais utilizados na diversas aplicações da TRI. Este Modelo é definido por:

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (5.7)$$

com  $i = 1, 2, \dots, I$ , e  $j = 1, 2, \dots, n$ , onde:

$U_{ij}$  é uma variável dicotômica que assume os valores 1, quando o indivíduo  $j$  responde corretamente o item  $i$ , ou 0 quando o indivíduo  $j$  responde incorretamente.

$\theta_j$  representa a habilidade (traço latente) do  $j$ -ésimo indivíduo.

$P(U_{ij} = 1|\theta_j)$  é a probabilidade de um indivíduo  $j$  com habilidade  $\theta_j$  responder corretamente o item  $i$  e é chamada de Função de Resposta do Item – FRI.

$b_i$  é o parâmetro de dificuldade (ou de posição) do item  $i$ , medido na mesma escala da habilidade.

$a_i$  é o parâmetro de discriminação (ou de inclinação) do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item — CCI no ponto  $b_i$ .

$c_i$  é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item  $i$  (muitas vezes referido como a probabilidade de acerto casual).

$e$  é a base dos logaritmos neperianos

$D$  é um fator de escala. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

Também nesse modelo não são considerados os valores negativos do parâmetro  $a_i$ .

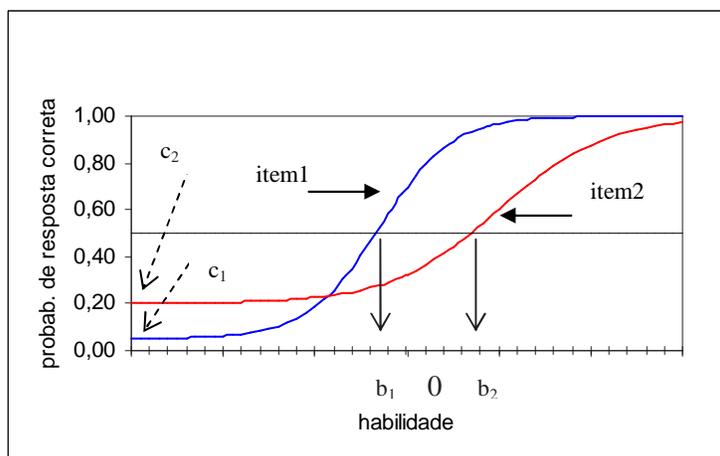


Figura 5.5 Curva Característica de dois Itens típicos do ML3

A Figura (5.5) mostra as CCIs de dois itens que diferem em três parâmetros: dificuldade (representada por  $b_1$  e  $b_2$ ), discriminação (que são as inclinações das CCIs dos itens 1 e 2 nos seus respectivos pontos de inflexão) e o chute (que é a probabilidade de sucesso por acaso). Observe que a CCI para o item 1 tem inclinação mais acentuada do que a CCI para o item 2, neste caso diz-se que o item 1 é melhor para discriminar do que o item 2. Analisando a CCI do item 2 nota-se que testandos com níveis baixos de habilidade são capazes de responder corretamente o item 2. E do ponto de vista da mensuração, um item com uma CCI semelhante a do item 2 pode ser menos eficiente do que o item 1.

### Função de Informação do Item e Função de Informação do Teste

Uma medida bastante utilizada em conjunto com a CCI é a *função de informação do item* ( $FII$ ). Ela permite analisar quanto um item (ou teste) contém de informação para a medida de habilidade.

Para o modelo logístico de 3 parâmetros, a equação pode ser escrita como:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[ \frac{P_i(\theta) - c_i}{1 - c_i} \right]^2$$

A função de informação fornecida pelo teste é definida como a soma das informações fornecidas por cada item que participa de sua composição. A função de informação de um item é dada por:

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

Para maiores detalhes sobre funções de informações, veja (Andrade, Tavares & Valle, 2000)

Na Tabela 5.1 são mostrados os parâmetros  $a$ ,  $b$  e  $c$  de quatro itens cujas CCIs e FIIs são apresentadas na Figura (5.6), sendo que as FIIs estão traçadas em linhas pontilhadas.

Tabela 5.1 *Tabela com quatro itens com seus respectivos parâmetros*

Item	$a$	$b$	$c$
1	1,40	-0,30	0,20
2	1,40	1,10	0,20
3	0,90	-0,30	0,20
4	0,90	1,10	0,20

Para cada item está associado um intervalo em torno do parâmetro  $b$  no qual o item tem maior poder de discriminação.

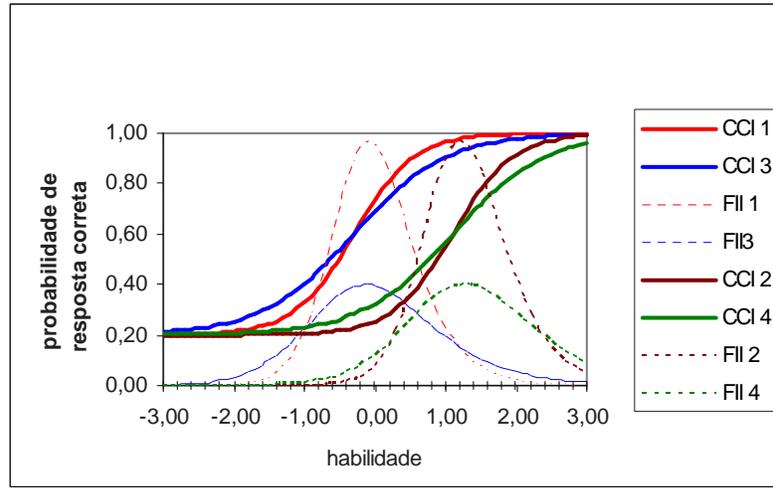


Figura 5.6 CCI e FII dos Itens 1, 2, 3 e 4 (da Tabela 5.1)

### 5.3 Modelos de Múltiplos Grupos

É bastante comum a situação de duas ou mais populações estarem envolvidas na análise. É o caso de grupos formados por diferentes graus de escolaridade, região, sexo, tipo de escola, etc. Para fazer a comparação dos grupos é preciso que tenham itens comuns nos testes aplicados à estes grupos. Em 1997, Bock & Zimowski propuseram um modelo onde considera-se que há  $K$  populações independentes em estudo e é feita uma análise conjunta das respostas amostrais dessas populações. Considera-se que a distribuição da habilidade dos indivíduos da população  $k$  segue uma determinada distribuição com vetor de parâmetros  $\boldsymbol{\eta}_k$ . Frequentemente adota-se a distribuição Normal com  $\boldsymbol{\eta}_k = (\mu_k, \sigma_k^2)'$ , sendo que estes parâmetros representam, respectivamente, a média e a variância das habilidades da população  $k$ ,  $k = 1, \dots, K$ . A função mais utilizada é

$$P(u_{kji} = 1 | \theta_{kj}) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_{kj} - b_i)}},$$

onde  $u_{kji}$  é a resposta (binária) ao item  $i$  oriunda do  $j$ -ésimo indivíduo do grupo  $k$ , e  $\theta_{kj}$  a habilidade do  $j$ -ésimo indivíduo do grupo  $k$ . (Por grupo  $k$  entenderemos a amostra relativa à população  $k$ .)

Para a construção do modelo são necessárias as seguintes suposições: independência local

e as respostas oriundas de indivíduos diferentes serão independentes. Maiores detalhes sobre este assunto, ver Andrade, Tavares & Valle (2000)

## 5.4 Métodos de Estimação

Na TRI, a probabilidade de resposta correta  $[P_i(\theta)]$  depende da habilidade do respondente e dos parâmetros dos itens. O objetivo da TRI é estabelecer modelos matemáticos que permitam estimar esses parâmetros e essas habilidades. O problema da estimação pode ser dividido em três casos:

Caso 1: já são conhecidas as habilidades dos respondentes e deseja-se estimar os parâmetros dos itens,

Caso 2: já são conhecidos os parâmetros dos itens e procura-se estimar as habilidades dos respondentes (ou seja, os itens já estão calibrados), e

Caso 3: deseja-se estimar, simultaneamente, os parâmetros dos itens e as habilidades dos respondentes.

Em qualquer um dos casos citados acima, geralmente a estimação é feita pelo Método da Máxima Verossimilhança através da aplicação de algum processo iterativo, como o algoritmo *Newton-Raphson* ou “*Scoring*” de Fisher. Alguns procedimentos bayesianos também são aplicados com bastante frequência (ver Mislevy, 1986a).

No caso em que desejamos estimar tanto os parâmetros dos itens, quanto as habilidades dos respondentes, dispõe-se usualmente de dois métodos: *Estimação Conjunta* (parâmetros dos itens e proficiências dos respondentes), e *Máxima Verossimilhança Marginal* (primeiramente a estimação dos parâmetros dos itens e depois das proficiências). Também são aplicados com bastante frequência alguns procedimentos bayesianos, como, por exemplo, a *Estimação pela média da distribuição posteriori*.

Para o desenvolvimento do modelo são necessárias as suposições: as respostas oriundas de indivíduos diferentes são independentes (S1), e os itens são respondidos de forma independente por cada indivíduo (Independência Local), fixada sua habilidade (S2). Serão usadas as seguintes notações:

$\theta_j$  representa a habilidade ou proficiência do respondente;

$U_{ji}$  é a variável aleatória que representa a resposta (binária) do indivíduo  $j$  ao item  $i$ , com

$$U_{ji} = \begin{cases} 1, & \text{resposta correta,} \\ 0, & \text{resposta incorreta.} \end{cases}$$

$\mathbf{U}_j = (U_{j1}, U_{j2}, \dots, U_{jI})$  o vetor aleatório de respostas do indivíduo  $j$  e

$\mathbf{U}_{..} = (\mathbf{U}_1, \dots, \mathbf{U}_n)$  o conjunto integral de respostas.

De forma similar, representaremos as observações por  $u_{ji}$ ,  $\mathbf{u}_j$  e  $\mathbf{u}_{..}$  e, ainda,

$\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$  representará o vetor de habilidades dos  $n$  indivíduos;

$\boldsymbol{\zeta}_i = (a_i, b_i, c_i)$  o vetor de parâmetro do item  $i$ ,  $i=1, \dots, I$ , e

$\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_I)$  o conjunto de parâmetros dos itens.

#### 5.4.1 Caso 1: já são conhecidas as habilidades dos respondentes e deseja-se estimar os parâmetros dos itens.

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), a verossimilhança  $L(\boldsymbol{\zeta}) = P(\mathbf{U}_{..} = \mathbf{u}_{..} | \boldsymbol{\theta}, \boldsymbol{\zeta})$ , pode ser escrita como:

$$\begin{aligned} L(\boldsymbol{\zeta}) &= \prod_{j=1}^n P(\mathbf{U}_j = \mathbf{u}_j | \theta_j, \boldsymbol{\zeta}) \\ &= \prod_{j=1}^n \prod_{i=1}^I P(U_{ji} = u_{ji} | \theta_j, \boldsymbol{\zeta}_i), \end{aligned}$$

onde na última igualdade usamos que a distribuição de  $U_{ji}$  só depende de  $\boldsymbol{\zeta}$  através de  $\boldsymbol{\zeta}_i$ .

Usando a notação  $P_{ji} = P(U_{ji} = 1 | \theta_j, \boldsymbol{\zeta}_i)$  e  $Q_{ji} = 1 - P_{ji}$ , temos que

$$\begin{aligned} P(U_{ji} = u_{ji} | \theta_j, \boldsymbol{\zeta}_i) &= P(U_{ji} = 1 | \theta_j, \boldsymbol{\zeta}_i)^{u_{ji}} P(U_{ji} = 0 | \theta_j, \boldsymbol{\zeta}_i)^{1-u_{ji}} \\ &= P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}. \end{aligned} \tag{5.8}$$

Portanto,

$$L(\zeta) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}. \quad (5.9)$$

Logaritmando-se ambos membros de (5.9)

$$\log L(\zeta) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}. \quad (5.10)$$

Os Estimadores de Máxima verossimilhança (EMV) de  $\zeta_i$ ,  $i = 1, \dots, I$ , são os valores que maximizam a equação (5.10). Em Andrade, Tavares & Valle são deduzidas as seguintes equações de estimação para os parâmetros dos itens:

$$a_i : \quad D(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji})(\theta_j - b_i) W_{ji} = 0, \quad (5.11)$$

$$b_i : \quad -Da_i(1 - c_i) \sum_{j=1}^n (u_{ji} - P_{ji}) W_{ji} = 0, \quad (5.12)$$

$$c_i : \quad \sum_{j=1}^n (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} = 0. \quad (5.13)$$

onde

$$W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}, \quad (5.14)$$

$$P_{ji}^* = \{1 + e^{-Da_i(\theta_j - b_i)}\}^{-1} \quad \text{e} \quad Q_{ji}^* = 1 - P_{ji}^*. \quad (5.15)$$

As equações (5.11), (5.12) e (5.13), não possuem solução explícita e por isso necessita-se de algum método iterativo para a obtenção das estimativas de máxima verossimilhança dos parâmetros dos itens.

#### 5.4.2 Caso 2: já são conhecidos os parâmetros dos itens e procura-se estimar as habilidades dos respondentes

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), podemos escrever a log-verossimilhança como em (5.10), agora como função de  $\theta$  e não de  $\zeta$ , ou seja,

$$\log L(\boldsymbol{\theta}) = \sum_{j=1}^n \sum_{i=1}^I \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}. \quad (5.16)$$

O EMV de  $\theta_j$  é o valor que maximiza a verossimilhança, ou equivalentemente, é a solução da equação

$$\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} = 0, \quad j = 1, \dots, n. \quad (5.17)$$

Notemos, de (5.16), que

$$\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} = \sum_{i=1}^I \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left( \frac{\partial P_{ji}}{\partial \theta_j} \right), \quad (5.18)$$

Como

$$\frac{\partial P_{ji}}{\partial \theta_j} = D a_i (1 - c_i) P_{ji}^* Q_{ji}^*, \quad (5.19)$$

obtêm-se

$$\begin{aligned} \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} &= \sum_{i=1}^I \left\{ (u_{ji} - P_{ji}) D a_i (1 - c_i) P_{ji}^* Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\ &= D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji}. \end{aligned} \quad (5.20)$$

Segue então que a equação de estimação (5.17) para  $\theta_j$ ,  $j = 1, \dots, n$ , é

$$\theta_j : \quad D \sum_{i=1}^I a_i (1 - c_i) (u_{ji} - P_{ji}) W_{ji} = 0. \quad (5.21)$$

Esta equação também não apresenta solução explícita para  $\theta_j$  e, por isso, precisa-se de algum método iterativo para obter as estimativas desejadas.

### 5.4.3 Caso 3: deseja-se estimar, simultaneamente, os parâmetros dos itens e as habilidades dos respondentes.

Para esta situação, a estimativa dos parâmetros dos itens e das habilidades dos respondente é feita conjuntamente, porém o método da estimação conjunta é muito trabalhoso computacionalmente. Ocorre a falta de identificabilidade, havendo necessidade de definir-se uma métrica, sendo que em muitos casos as habilidades e os parâmetros dos itens são estimados na métrica  $(\mu, \sigma)$ , onde  $\mu=0$  e  $\sigma=1$ . Uma alternativa para o método da máxima verossimilhança conjunta é o *Método da Máxima Verossimilhança Marginal*

### 5.4.4 Máxima verossimilhança marginal

A proposta básica do método da Máxima Verossimilhança Marginal é fazer a estimação em duas etapas: primeiro os parâmetros dos itens e, posteriormente, as habilidades dos respondentes. Esse método, proposto por Bock & Lieberman (1970), apresenta algumas vantagens em relação ao método da Máxima Verossimilhança Conjunta. Ele consiste no uso de um artifício que faz com que a verossimilhança torne-se função apenas dos parâmetros dos itens. Para Andersen (1980), se considerarmos uma população  $\Pi$  composta por  $n$  indivíduos com habilidades  $\theta_j$ ,  $j = 1, \dots, n$ , e construirmos a distribuição de frequência acumulada  $G(\theta) = (\text{número de } j : \theta_j \leq \theta) / n$ , então, se  $n$  for suficientemente grande os  $\theta_j$  estarão bastante próximos de forma que  $G(\theta)$  pode ser aproximada por uma distribuição contínua. A densidade  $g(\theta)$ , relativa à  $G(\theta)$ , pode realmente ser considerada a função densidade para  $\theta$  no experimento de retirar um indivíduo ao acaso da população  $\Pi$  e observar seu parâmetro  $\theta$ . Ao atribuímos uma distribuição de probabilidade para  $\theta$  *não estamos aplicando nenhum argumento bayesiano*. A distribuição de  $\theta$  realmente existe, no sentido explicado acima, como a densidade relativa à distribuição  $G(\theta)$ , (Andrade, Tavares & Valle, 2000). Desse modo, para eliminar as habilidades na verossimilhança, o artifício consiste em marginalizar a verossimilhança integrando-a com relação à distribuição da habilidade.

De forma geral, consideremos que as habilidades,  $\theta_j$ ,  $j = 1, \dots, n$ , são realizações de uma variável aleatória  $\theta$  com distribuição contínua e função densidade de probabilidade (*fdp*)  $g(\theta|\boldsymbol{\eta})$ , duplamente diferenciável, com as componentes de  $\boldsymbol{\eta}$  conhecidas e finitas. Para o caso em que  $\theta$  tem distribuição Normal, temos  $\boldsymbol{\eta} = (\mu, \sigma^2)$ , onde  $\mu$  é a média e

$\sigma^2$  a variância das habilidades dos indivíduos de  $\Pi$ . Portanto, se desejarmos que os itens sejam estimados na métrica (0,1), deveremos adotar  $\mu = 0$  e  $\sigma = 1$ . Bock & Lieberman assumiram que, de acordo com as notações acima, a probabilidade marginal de  $\mathbf{U}_j$  é dada por

$$\begin{aligned} P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \int_{\mathbb{R}} P(\mathbf{u}_j|\theta, \boldsymbol{\zeta}, \boldsymbol{\eta})g(\theta|\boldsymbol{\eta})d\theta \\ &= \int_{\mathbb{R}} P(\mathbf{u}_j|\theta, \boldsymbol{\zeta})g(\theta|\boldsymbol{\eta})d\theta, \end{aligned} \quad (5.22)$$

onde na última igualdade usamos que a distribuição de  $\mathbf{U}_j$  não é função de  $\boldsymbol{\eta}$ . Pela independência entre as respostas de diferentes indivíduos, podemos escrever a verossimilhança como:

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = P(\mathbf{u}_..|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{j=1}^n P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}). \quad (5.23)$$

Usando a abordagem de *Padrões de Resposta* (Andrade, Tavares e Valle, 2000), irá facilitar os cálculos computacionais. Como temos  $I$  itens no total, com 2 possíveis respostas para cada item (0 ou 1), há  $S = 2^I$  possíveis respostas (padrões de resposta). Quando o número de indivíduos é grande com relação ao número de itens, pode haver vantagens computacionais em trabalhar com o número de ocorrências dos diferentes padrões de resposta. Neste sentido, daqui em diante vamos trabalhar considerando este raciocínio. O índice  $j$  não mais representará um indivíduo, mas sim um padrão de resposta (Andrade, Tavares & Valle, 2000). Seja  $r_j$  o número de ocorrências distintas do padrão de resposta  $j$ , do tipo  $(u_{j1}, \dots, u_{jn})$ , e ainda  $s \leq \min(n, S)$  o número de padrões de resposta com  $r_j > 0$ . Segue disso que

$$\sum_{j=1}^s r_j = n. \quad (5.24)$$

Pela independência entre as respostas dos diferentes indivíduos, temos que os dados seguem uma distribuição *Multinomial* (Andrade, Tavares e Valle, 2000), isto é,

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{n!}{\prod_{j=1}^s r_j!} \prod_{j=1}^s [P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta})]^{r_j}. \quad (5.25)$$

Igualando-se a derivada da log-verossimilhança, com relação a  $\zeta_i$ , ao vetor nulo, obtêm-se as equações de estimação para os parâmetros dos itens:  $a_i$ ,  $b_i$  e  $c_i$ , respectivamente,

$$a_i : D(1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(u_{ji} - P_i)(\theta - b_i)W_i] g_j^*(\theta) d\theta = 0, \quad (5.26)$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^s r_j \int_{\mathbb{R}} [(u_{ji} - P_i)W_i] g_j^*(\theta) d\theta = 0, \quad (5.27)$$

$$c_i : \sum_{j=1}^s r_j \int_{\mathbb{R}} \left[ (u_{ji} - P_i) \frac{W_i}{P_i^*} \right] g_j^*(\theta) d\theta = 0, \quad (5.28)$$

as quais não possuem solução explícita, sendo portanto necessário métodos iterativos para resolvê-las. As equações envolvem integrais que não apresentam soluções analíticas e devem ser resolvidas por algum método de quadratura. Na TRI costuma-se usar o método *Hermit-Gauss* (Veja Andrade, Tavares & Valle, 2000). As equações de estimação em forma de quadratura para os parâmetros:  $a_i$ ,  $b_i$  e  $c_i$  são, respectivamente,

$$a_i : D(1 - c_i) \sum_{j=1}^s \sum_{k=1}^q r_j [(u_{ji} - P_{ki})(\bar{\theta}_k - b_i)W_{ki}] g_j^*(\bar{\theta}_k). \quad (5.29)$$

$$b_i : -Da_i(1 - c_i) \sum_{j=1}^s \sum_{k=1}^q r_j [(u_{ji} - P_{ki})W_{ki}] g_j^*(\bar{\theta}_k) = 0, \quad (5.30)$$

$$c_i : \sum_{j=1}^s \sum_{k=1}^q r_j \left[ (u_{ji} - P_{ki}) \frac{W_{ki}}{P_{ki}^*} \right] g_j^*(\bar{\theta}_k) = 0. \quad (5.31)$$

Essas equações devem ser resolvidas por algum processo iterativo, tipo Newton-Raphson. Nota-se, no entanto, que nestes processos pode haver necessidade da derivada segunda, e estas derivadas segundas entre (parâmetros de) itens diferentes são não nulas, o que dificulta o processo. Se o número de itens for grande ocorrerá limitação na estimação dos parâmetros e, para solucionar este problema, Bock & Aitkin (1981), reformularam a abordagem proposta por Bock & Lieberman. Eles apresentaram uma proposta que inclui a construção de variáveis não observáveis, representadas por  $f$  e  $r$  a seguir. Elas concentram a dependência entre as derivadas segundas citadas acima, de forma que nesta proposta, fixados  $f$  e  $r$ , as derivadas segundas entre itens distintos passam a não existir, facilitando em muito o processo de estimação. Por Bock & Aitkin (1981) temos as seguintes equações de estimação em forma de quadratura para os parâmetros:  $a_i$ ,  $b_i$  e  $c_i$ :

$$a_i : D(1 - c_i) \sum_{k=1}^q (\bar{\theta}_k - b_i) [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0, \quad (5.32)$$

$$b_i : -D a_i (1 - c_i) \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0, \quad (5.33)$$

$$c_i : \sum_{k=1}^q [r_{ki} - P_{ki} f_{ki}] \frac{W_{ki}}{P_{ki}^*} = 0. \quad (5.34)$$

onde

$$r_{ki} = \sum_{j=1}^s r_j u_{ji} g_{jk}^*, \quad f_k = \sum_{j=1}^s r_j g_{jk}^* \quad \text{e} \quad g_{jk}^* = g_j^*(\bar{\theta}_k). \quad (5.35)$$

Essas equações são resolvidas aplicando-se o algoritmo EM (Dempster, Laird & Rubin, 1977) que é um processo iterativo realizado em dois passos: Esperança (E) e Maximização (M). Mais especificamente, os passos E e M são

**Passo E** Usar os pontos de quadratura  $\bar{\theta}_k$ , os pesos  $A_k$ ,  $k = 1, \dots, q$  e estimativas iniciais dos parâmetros dos itens,  $\hat{\zeta}_i$ ,  $i = 1, \dots, I$ , para gerar  $g_j^*(\bar{\theta}_k)$  e, posteriormente,  $r_{ki}$  e  $f_k$ ,  $i = 1, \dots, n$  e  $k = 1, \dots, q$ .

**Passo M** Com  $\mathbf{r} = (r_{ki})$  e  $\mathbf{f} = (f_k)$  obtidos no Passo E, resolver as equações de estimação para  $\zeta_i$ ,  $i = 1, \dots, n$ , usando o algoritmo Newton-Raphson ou “Scoring” de Fisher.

Estes passos compõem cada iteração do algoritmo EM, as quais serão repetidas até que algum critério de parada seja alcançado. Após a finalização do processo, os erros-padrão são obtidos com o uso da matriz de informação de Fisher. Para mais informações sobre o algoritmo EM veja Andrade, Tavares & Valle (2000).

Concluída a fase de estimação dos parâmetros dos itens, inicia-se a etapa de estimação das proficiências individuais. Nesta etapa considera-se que os parâmetros dos itens já são conhecidos, e são adotadas as estimativas usadas na primeira etapa. Na próxima seção aborda-se os principais métodos dessa estimação.

## 5.5 Estimação das proficiências

Basicamente, supôs-se que as proficiências da população estudada compunham uma determinada distribuição, tal como a normal, com densidade denotada por  $g(\theta)$ . Assim sendo, é plausível acreditar que a proficiência de um determinado indivíduo dessa população é dada pela distribuição da população, dado o conjunto de respostas do indivíduo. Ainda, adotando-se independência entre as respostas de diferentes indivíduos, podemos fazer as estimações em separado para cada indivíduo.

Nestes termos, a estimação da habilidade do indivíduo  $j$  é baseada na função

$$g_j^*(\theta) \equiv g(\theta|\mathbf{u}_j, \boldsymbol{\zeta}, \boldsymbol{\eta}) \propto P(\mathbf{u}_j|\theta_j, \boldsymbol{\zeta})g(\theta_j|\boldsymbol{\eta}). \quad (5.36)$$

Novamente, podemos adotar alguma característica de  $g_j^*(\theta_j)$  como estimador de  $\theta_j$ , sendo que as mais adotadas são a média e a moda. A seguir, trataremos unicamente da obtenção das habilidades pela média da posteriori, por ser o método mais aplicado na área de avaliação educacional, assim sendo o padrão nos estudos do Inep/MEC.

### 5.5.1 Estimação das Habilidades pela Média da Posteriori - EAP

O principal método de estimação utilizado atualmente tem por base a Máxima Verossimilhança Marginal (MVM). Supondo-se a independência entre as habilidades de diferentes indivíduos, é possível fazer a estimação das habilidades em separado para cada um. Vamos assumir que a distribuição a priori para a habilidade da população é Normal, com vetor de parâmetros  $\boldsymbol{\eta}=(\mu, \sigma^2)$  conhecidos, em geral  $\mu = 0$  e  $\sigma^2 = 1$ , que são o padrão quando da calibração dos itens.

A posteriori para a habilidade do indivíduo  $j$  é dada por (5.36) e a média da respectiva distribuição como estimador de  $\theta$  é o que denomina pelo estimador EAP. Como

$$g(\theta|\mathbf{u}_j, \boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{P(\mathbf{u}_j|\boldsymbol{\zeta}, \theta)g(\theta|\boldsymbol{\eta})}{P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta})}. \quad (5.37)$$

onde,

$$P(\mathbf{u}_j|\boldsymbol{\zeta}, \theta) = \prod_{j=1}^n P(\mathbf{u}_{ji}|\boldsymbol{\zeta}_i, \theta) = \prod_{j=1}^n P_{ji}(\theta)^{u_{ji}} [1 - P_{ji}(\theta)]^{1-u_{ji}} \quad (5.38)$$

e

$$P(\mathbf{u}_j|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \int_{\mathcal{R}} P(\mathbf{u}_j|\theta, \boldsymbol{\zeta})g(\theta|\boldsymbol{\eta})d\theta. \quad (5.39)$$

Segue que a esperança da posteriori é

$$\hat{\theta}_j \equiv E[\theta | \mathbf{u}_j, \boldsymbol{\zeta}, \boldsymbol{\eta}] = \frac{\int_{\mathbb{R}} \theta g(\theta | \boldsymbol{\eta}) P(\mathbf{u}_j | \theta, \boldsymbol{\zeta}) d\theta}{\int_{\mathbb{R}} g(\theta | \boldsymbol{\eta}) P(\mathbf{u}_j | \theta, \boldsymbol{\zeta}) d\theta}. \quad (5.40)$$

Esta forma de estimação tem a vantagem de ser calculada diretamente, não necessitando da aplicação de métodos iterativos. Além disso, as quantidades necessárias para o seu cálculo são um produto final da etapa de calibração. Por conta disso alguns autores (por exemplo, Mislevy & Stocking, 1989) recomendam esta escolha para a estimação das habilidades. Maiores detalhes em (Andrade, Tavares & Valle, 2000). Algumas vezes, em Educação, necessita-se trabalhar com mais de uma população, neste caso utiliza-se modelos de múltiplos grupos.

## 5.6 Sobre a escala

Depois de estimados os parâmetros dos itens e as proficiências dos alunos, em geral na escala padrão ( $\mu = 0, \sigma = 1$ ), representada apenas por  $(0,1)$ , os resultados são colocados em uma outra escala de desempenho pré-determinada. O Saeb, por exemplo, usa a escala  $(250,50)$ . Essa transformação é linear para as proficiências, de forma que

$$\theta^* = \sigma * \theta + \mu \quad (5.41)$$

será a proficiência na escala  $(\mu, \sigma)$ . No Saeb, quando de sua primeira aplicação em 1997, e como sempre ocorre, verificou-se que a média e o desvio padrão daquelas estimativas não davam exatamente 0 e 1, de forma que, para fins de garantia que as proficiências daquele ano fossem realmente 0 e 1, precisava haver uma certa padronização para  $(90,1)$  e, posteriormente, a transformação para  $(250,50)$ . Com isso, as constantes de transformação não foram exatamente 250 e 50, mas as proficiências médias daquele ano são exatamente 250 e 50.

O trabalho restante consiste em definir pedagogicamente o que cada valor, ou faixas, representam em termos de conhecimento, e quais fatores associados (intra-estrutura da escola, formação dos pais, nível sócio-econômico etc.) podem estar interferindo no valor desta proficiência.

Há vários softwares comerciais usados na estimação de parâmetros dos itens e proficiências.

De forma a dar uma idéia de como se apresentam os comandos básicos dos programas, apresentamos os softwares mais conhecidos e os comandos usados na estimação das proficiências do NPI.

## 5.7 Aspectos computacionais

As ferramentas matemáticas utilizadas no desenvolvimento da Teoria da Resposta ao Item são bastante complexas e por isso quem trabalha com essa Teoria depende de recursos computacionais consideráveis. Neste capítulo serão vistos primeiramente alguns exemplos de recursos computacionais e a seguir, um estudo sobre o programa Bilog-Mg.

### 5.7.1 Principais softwares

Para ajustar os modelos de TRI são utilizados programas computacionais como, por exemplo:

1. Bilog e Bilog-Mg

Esses programas são específicos para a análises via TRI de itens dicotômicos ou dicotomizados. O Bilog permite o ajuste do modelo da TRI para um único grupo de respondentes, enquanto que o Bilog-Mg permite a análise de mais de um grupo.

2. Parscale

Esse programa analisa itens não dicotômicos e tem implementado os modelos de Resposta Gradual e de Créditos Parciais, podendo fazer análises para mais de um grupo de respondentes.

3. Testfact

Este programa produz uma análise fatorial de verossimilhança completa para os itens de um teste. Ele oferece recursos importantes para a TRI, usados na verificação da dimensionalidade dos testes.

#### 4. Multilog

Faz o ajuste de modelos unidimensionais da TRI para dados dicotômicos ou politômicos.

Os modelos considerados são o logístico de 1, 2 e 3 parâmetros, de Resposta Nominal, de Crédito Parcial. Permite também o ajuste para múltiplos grupos.

#### 5. Iteman

Faz a análise clássica dos itens.

### 5.7.2 Bilog, Bilog-Mg

Esses programas executam a análise dos itens em três etapas, chamadas de *Fase 1*, *Fase 2* e *Fase 3*.

-A fase 1 é a fase de entrada e leitura dos dados. O usuário fornece ao programa a identificação de cada respondente com suas respectivas respostas (A,B,C,..., ou 0,1,2,3,...) ao teste e o gabarito. Caso seja apresentado ao programa as respostas corrigidas (1 como acerto e 0 como erro) não haverá necessidade de apresentar o gabarito. Os arquivos de saída fornecem algumas estatísticas descritivas e correlações ponto bisserial de interesse usada na Teoria Clássica dos Testes. Em casos de amostragem complexa o programa usa os pesos dos respondentes para calibração e estimação das proficiências.

-A fase 2 é a fase de calibração dos itens com obtenção das estimativas pontuais dos parâmetros dos itens e seus respectivos erros-padrão.

-A fase 3 é a fase de estimação das habilidades dos respondentes.

O Bilog e o Bilog-Mg fornecem gráficos contendo algumas informações como as curvas características e as curvas de informação de cada item e do teste. Tanto o Bilog quanto o Bilog-Mg têm implementado os modelos unidimensionais logísticos de 1, 2 e 3 parâmetros e a diferença entre eles é que o Bilog-Mg fornece a análise de mais de um grupo de respondentes enquanto que o Bilog analisa apenas um grupo de respondentes. A seguir mostra-se um exemplo da aplicação do Bilog-Mg em que são usados os seguintes comandos: >COMMENT, >GLOBAL, >SAVE, >LENGTH, etc, Neste exemplo, os números que aparecem no comando SLOpe são os parâmetros de discriminação dos itens. Os parâmetros de dificuldade estão discriminados no comando THReshld, e em GUEss estão

os parâmetros correspondentes ao “chute”.

Exemplo de uma aplicação do BilogMg:

```
>COMMENT
>GLOBAL DFName = 'PORTNPI08.PRN',NPArm = 3,SAVe;
>SAVE PArM = 'PORTNPI08.par', SCOrE = 'PORTNPI08.sco';
>LENGTH NITems = (39);
>INPUT NTOTal = 39,SAMple = 1000000,NALt = 4,NIDchar = 6,
      KFName = 'GABARITOP08.TXT';
>ITEMS INAmes = (Item01(1)Item39);
>TEST1 TNAmE = 'Port08',INUmber = (1(1)39),
      SLOpe = (1.7380, 1.3310, 1.5040, 1.3310, 2.4050, 1.6730,
              1.3200, 1.0430, 1.1470, 2.6220, 1.6610, 1.8460,1.5610,
              1.6520, 2.3400, 2.2150, 2.0070, 2.0420, 1.1340, 0.9300,
              1.0980, 1.2590, 0.9780, 0.9680, 1.1720, 1.8340, 1.6820,
              1.1980, 1.3760, 1.0500, 2.1600, 2.3080, 0.9770, 0.9870,
              1.1430, 2.6410, 1.2850, 1.3580, 1.4010),
      THReshld = (-0.3210, 0.0710,-0.0050, 0.8900, 0.5540,
                 -0.2550, 0.0880, 0.1480, 0.5330, -0.6180,-0.4000,
                 -1.0480, 0.1320,-0.0980, 0.2000, -0.1430, 0.4600,
                 -0.0480,-0.5490,-2.3620,-0.7080, -0.8710,-1.3910,
                 1.2360,-0.2660, 1.0480,-1.3550, -1.3010,-0.6350,
                 -1.1100, 0.5220,-0.1550,-0.7420, -0.7790, 0.5950,
                 0.7210, 0.1560, 0.5880,-0.1620),
      GUEss = (0.2530, 0.1790, 0.2350, 0.2370, 0.2760, 0.2070,
              0.0920, 0.2020, 0.2800, 0.2400, 0.2790, 0.0430, 0.1980,
              0.1780, 0.1920, 0.1780, 0.2710, 0.2580, 0.2050, 0.0210,
              0.1260, 0.0260, 0.0780, 0.2560, 0.1510, 0.2790, 0.0850,
              0.0840, 0.1880, 0.1120, 0.1810, 0.2440, 0.1300, 0.2210,
              0.1140, 0.1800, 0.2700, 0.2520, 0.0920);

(6A1,39A1)
>CALIB SELEct = (0),
      ACCel = 1.0000;
>SCORE NQPt = (20),
      NOPrint;
```

É na fase 3 que podemos, dentre outras coisas, definir a escala de proficiências para que as estimativas sejam transformadas pelo próprio programa. Este software é bastante customizável, e tornou-se o padrão na área de avaliação educacional, particularmente nos procedimentos realizados para o Inep. Maiores informações sobre o mesmo podem ser obtidas em [www.ssicentral.com](http://www.ssicentral.com).

Enfim, visto o processo de estimação e mostrado os passos do software básico da área, caminhamos rumo à aplicação prevista no trabalho. Antes, porém, será feita uma breve descrição das escala de proficiências adotado pelo Saeb para que os reultados a serem apresentados no capítulo posterior possam ser melhor interpretados.

## ESCALAS DE DESEMPENHO

---

### 6.1 Introdução

Quando compara-se, por exemplo, a habilidade média em matemática entre os alunos da 5.<sup>a</sup> e da 6.<sup>a</sup> série do ensino fundamental e observa-se um ganho “ $\Delta$ ”, faz-se apenas uma comparação quantitativa; mas qualitativamente o que os alunos da série mais adiantada *sabem a mais* que os alunos da série mais atrasada em termos de conteúdo? Para buscar uma interpretação qualitativa dos valores obtidos é que são construídas *Escala de Desempenho ou Habilidades*, utilizando-se para isso uma metodologia matemática e todo um trabalho de interpretação pedagógica.

A escala de desempenho é uma escala *arbitrária* onde o que interessa na verdade são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude. Através delas pode-se conhecer o percentual de alunos que já construíram as competências e habilidades desejáveis para cada uma das séries avaliadas, quantos estão ainda em processo de construção, quantos estão abaixo do nível que seria desejável para a série e quantos estão acima do nível esperado. As escalas de desempenho ordenam o desempenho dos alunos do menor para o maior em um continuum e elas são cumulativas, ou seja, o que os alunos sabem, compreendem e são capazes de fazer quando seu desempenho situa-se em um nível da escala, são capazes também de demonstrar as habilidades descritas no(s) nível(eis) anterior(es) dessa escala.

As escalas de habilidade são definidas por níveis denominados de *níveis âncoras* que são caracterizados por *itens âncoras*. Exemplo de níveis âncoras são: Nível 1 - Habilidades entre 125 e 150; Nível 2 - Habilidades entre 150 e 175, etc. Os itens âncoras indicarão o conhecimento absorvido por indivíduos nestas faixas de habilidade.

A definição de itens âncoras é a seguinte: Considere dois níveis âncora consecutivos Y

e  $Z$  com  $Y$  menor que  $Z$ . Dizemos que um determinado item é âncora para o nível  $Z$  se, e somente se, as três condições abaixo forem satisfeitas simultaneamente :

a)  $P(U=1 | \theta = Z) \geq 0,65$  e

b)  $P(U=1 | \theta = Y) < 0,50$  e

c)  $P(U=1 | \theta = Z) - P(U=1 | \theta = Y) \geq 0,30$

Para elucidar essa definição considera-se um item que está medindo a mesma habilidade dos respondentes. Chamaremos esse item de *item A*. Se mais de 65% dos respondentes do nível  $Z$  acertarem esse item e menos de 50% errarem esse mesmo item, e além disso, se a diferença entre essas duas proporções for maior do que 30% então o *item A* é um item âncora.

## 6.2 A Escala do Saeb

A escala do Saeb é única para cada disciplina e permite apresentar, em uma mesma métrica, os resultados de desempenho dos estudantes de todas as séries (4.<sup>a</sup> e 8.<sup>a</sup> série do Ensino Fundamental e 3.<sup>a</sup> série do Ensino Médio) e anos de aplicação das provas (1995, 1997, 1999, 2001, 2003 e 2005). As escalas usadas pelo Saeb (de Língua Portuguesa e de Matemática) variam predominantemente de 0 a 500 pontos. Nelas o desempenho está ordenado de maneira crescente e cumulativa (o nível 500 abrange todos os anteriores), a estimativa de desempenho obtida para os alunos ou grupos de alunos em cada uma das três séries avaliadas indica, portanto, o lugar que ocupam na escala.

A escala do Saeb para Língua Portuguesa é única e possui 8 níveis de desempenho: Nível 1 (125 a 150), Nível 2 (150 a 175), Nível 3 (175 a 200), Nível 4 (200 a 250), Nível 5 (250 a 300), Nível 6 (300 a 350), Nível 7 (350 a 375) e Nível 8 (375 e acima). Para a 4.<sup>a</sup> série do Ensino Fundamental, são considerados os cinco primeiros níveis; para a 8.<sup>a</sup> série são considerados os níveis 2, 3, 4, 5, 6 e 7. Para a 3.<sup>a</sup> série do Ensino Médio considera-se os níveis 5, 6, 7 e 8. A escala do Saeb para a Matemática também é única e é formada por dez níveis de desempenho: Nível 1 (125 a 150), Nível 2 (150 a 175), Nível 3 (175 a 200), Nível 4 (200 a 250), Nível 5 (250 a 300), Nível 6 (300 a 350), Nível 7 (350 a 375), Nível 8 (375 a 400), Nível 9 (400 a 425) e Nível 10 (425 e acima). Para a 4.<sup>a</sup> série do Ensino

Fundamental, são considerados os níveis 1, 2, 3, 4, 5 e 6. Para a 8ª série são considerados os níveis 4, 5, 6, 7, e 8. Para a 3ª série do Ensino Médio são considerados os níveis 5, 6, 7, 8, 9, e 10.

Tanto para Língua Portuguesa como para Matemática, os níveis são cumulativos, isto é, os alunos posicionados em um nível dominam as habilidades descritas no(s) nível(is) anterior(es) da escala. Em cada nível é apresentado o percentual de alunos que dominam as habilidades nele descritas para cada série. Assim pode-se verificar que percentual de alunos já construiu as competências e habilidades desejáveis para cada uma das séries avaliadas, quantos ainda estão em processo de construção, quantos estão abaixo do nível que seria desejável para a série e quantos estão acima do nível que seria esperado.

No próximo capítulo são apresentados alguns recursos computacionais usados na Teoria da Resposta ao Item.

## **Resultado da Pesquisa: Posicionamento do NPI Escala do Saeb**

---

### **7.1 Introdução**

O Núcleo Pedagógico Integrado (NPI), hoje chamado de Escola de Aplicação da Universidade Federal do Pará, funciona com as seguintes modalidades de ensino: Educação Infantil, Ensino Fundamental, Ensino Médio e Ensino de Jovens e Adultos. A Escola possui vários serviços para o alunado, como: Biblioteca, Laboratório de Informática, Laboratório de Ciências, Laboratório de Línguas Estrangeiras, Reprografia, Sala de Multimídia etc. e quase todos os professores possuem pós-graduação.

Para realizar este trabalho, aplicou-se provas de Língua Portuguesa e de Matemática, aos alunos das 4.<sup>a</sup> e 8.<sup>a</sup> séries do Ensino Fundamental, juntamente com questionários para os alunos, para os professores das séries avaliadas e para o diretor da escola. As provas continham 39 itens de múltipla escolha já calibrados pelo Inep, por isso foi feita apenas a estimativa das proficiências dos alunos. Participaram desta pesquisa 132 alunos da 4<sup>a</sup> série e 153 da 8<sup>a</sup>. Os resultados de desempenho dos alunos estão dispostos em duas formas: as médias gerais de proficiência e o desempenho dos alunos por estágios de competências. A análise dos questionários ficará para um trabalho futuro.

### **7.2 Resultado da Pesquisa**

#### **7.2.1 Médias Gerais de Proficiência**

Estimadas as habilidades dos alunos pelo método EAP (média da posteriori), foram determinadas as habilidades médias dos alunos por disciplina e série. São essas habilidades médias que são colocadas na escala do Saeb. Analisaremos separadamente cada uma das disciplinas avaliadas.

### Língua Portuguesa

A proficiência média em Língua Portuguesa da 4ª série foi 189,64 pontos, com desvio-padrão igual a 40,36. Com esse resultado, o NPI ocupa o nível 3 da Escala de Desempenho (em Língua Portuguesa) do Saeb, com um total de pontos maior que a média nacional da 4ª série, em 2005 (ver Relatório Saeb, 2002 - Língua Portuguesa). Nesse nível os alunos são capazes de localizar informações explícitas em textos mais complexos, mais longos e não narrativos (textos publicitários), inferir o sentido de palavra ou expressão em textos narrativos, etc. .

A proficiência média dos alunos da 8ª série em Língua Portuguesa foi 242,26 pontos, com desvio-padrão igual a 50,91. Esse resultado coloca a escola em questão, no nível 4 da Escala de Desempenho do Saeb. Neste nível, o aluno identifica o tema em textos poéticos, estabelece relações de continuidade relacionando uma palavra de sentido mais genérico a outra de sentido mais específico, etc.

### Matemática

A proficiência média em Matemática dos alunos da 4ª série foi 201,89 pontos, com desvio padrão igual a 33,07. Com esse resultado, o NPI ocupa o nível 4 da escala do Saeb. Estar no Nível 4 de desempenho significa que os alunos já desenvolveram certas habilidades, como efetuar as quatro operações aritméticas, que são importante para a resolução de problemas do cotidiano. O estudante que não dominar esse pré-requisito, terá consideráveis dificuldades na sua trajetória escolar. O estudante da 4ª série do Ensino Fundamental deve dominar os seis primeiros níveis da escala única de matemática que, como já foi visto no Capítulo anterior, vai de 1 até o Nível 6, para a 4ª série, entretanto apenas uma pequena quantidade de alunos conseguiu chegar no Nível 6. Os alunos no nível 4 são capazes de distinguir sólidos que têm superfície arredondada. Lêem horas e minutos em relógios de ponteiros, realizam conversões entre medidas de tempo (hora, dia e semana), etc.

Quanto aos alunos da 8ª série, estes conseguiram 265,63 pontos, com desvio-padrão de 37,92, ficando situados no nível 5 da escala do Saeb. Neste nível os alunos calculam

volume através de contagem de blocos, reconhecem gráficos de colunas referentes a dados apresentados de forma textual, etc.

No Apêndices: C, D, E e F, baseado na escala do Saeb, pode-se ver o que os alunos da 4<sup>a</sup> e da 8<sup>a</sup> séries do NPI já dominam em Matemática e em Língua Portuguesa.

As Figuras 7.1 e 7.2 mostram a posição das médias de proficiência em Língua Portuguesa e em Matemática dos alunos do NPI em relação aos alunos brasileiro de um modo geral cujo resultado estão no Apêndice B.

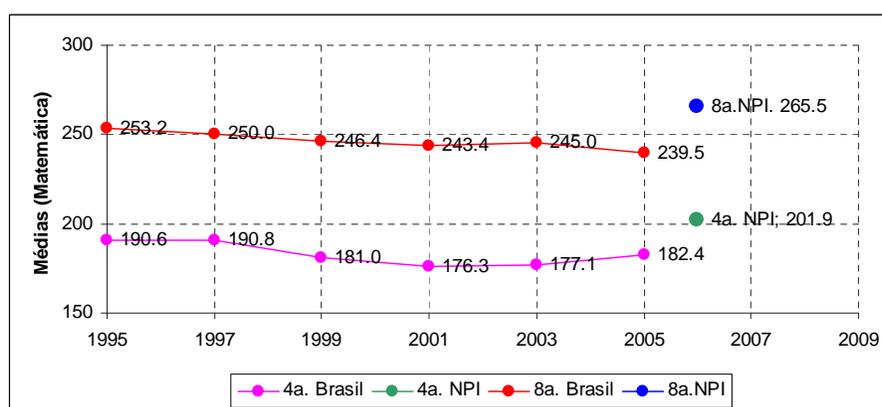


Figura 7.1 Brasil - Proficiências em Matemática (Saeb: 1995 - 2005 e NPI: 2006)

Fonte: INEP & Autora

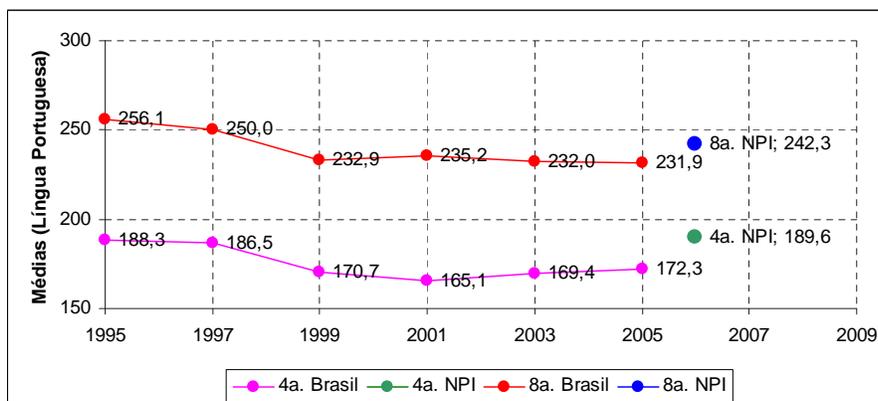


Figura 7.2 Brasil - Proficiências em Língua Portuguesa (Saeb: 1995 - 2005 e NPI - 2006)

Fonte: INEP & Autora

Na Tabela 7.1 e nas Figuras de 7.1 à 7.4 estão as médias de proficiências em Língua Portuguesa e Matemática: Saeb 2005 (Brasil, Região Norte e Estado do Pará) e NPI, juntamente com os respectivos desvios-padrão.

Tabela 7.1 Proficiências do Saeb - 2005 - (Brasil, Região Norte, Estado do Pará), e do NPI - 2006

Disciplina	Séries	Brasil 2005	Norte 2005	Pará 2005	NPI 2006
Matemática	4 <sup>a</sup>	182,37 (47,08)*	166,97 (37,13)	163,50 (35,04)	201,89 (33,07)
Matemática	8 <sup>a</sup>	239,52 (50,92)	226,80 (43,67)	230,59 (41,95)	265,63 (37,92)
Língua Portuguesa	4 <sup>a</sup>	172,31 (44,57)	160,50 (38,91)	158,85 (39,17)	189,65 (40,36)
Língua Portuguesa	8 <sup>a</sup>	231,81 (49,48)	225,84 (43,34)	231,55 (42,64)	242,26 (50,91)

\* : Desvio Padrão

Fonte Adaptado: Inep - NPI

Com os dados da Tabela 7.1 foram construídas as Figura 7.3, Figura 7.4, Figura 7.5 e Figura 7.6.

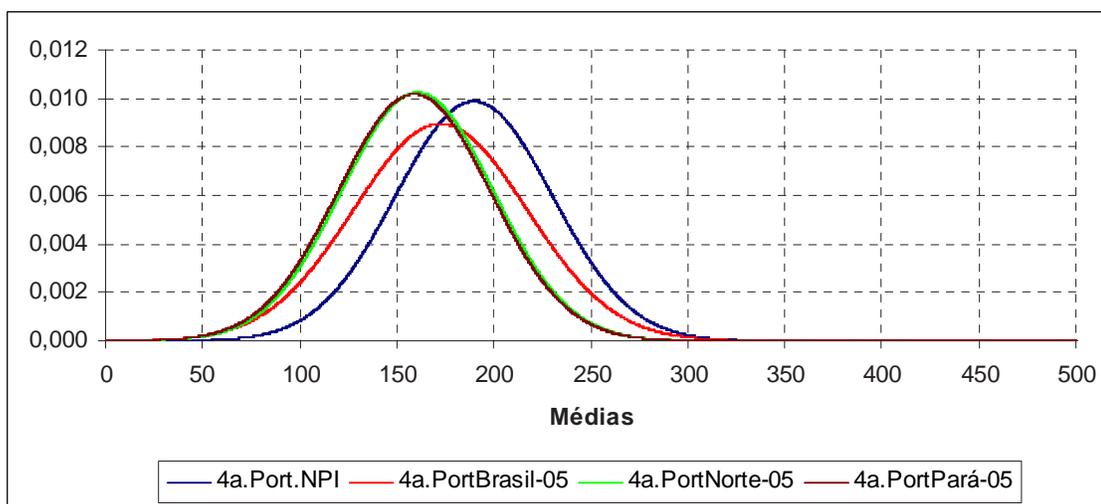


Figura 7.3 Proficiências Médias em Língua Portuguesa - 4ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará.

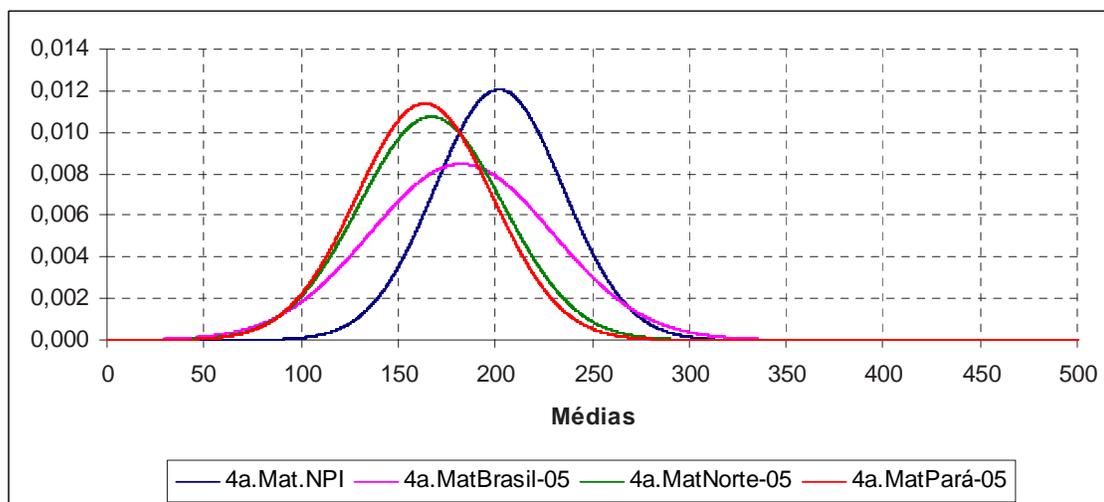


Figura 7.4 Proficiências Médias em Matemática - 4ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará.

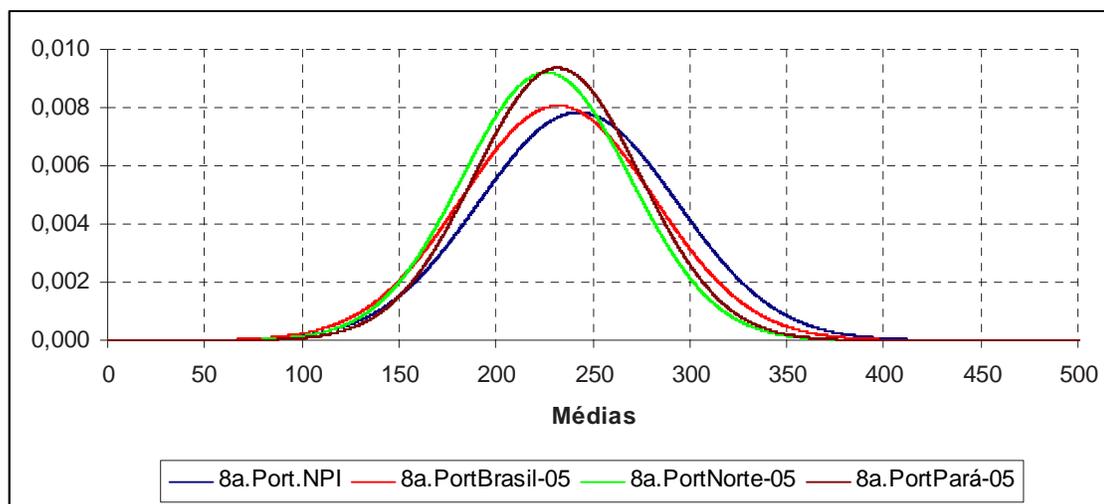


Figura 7.5 *Proficiências Médias em Língua Portuguesa - 8ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará.*

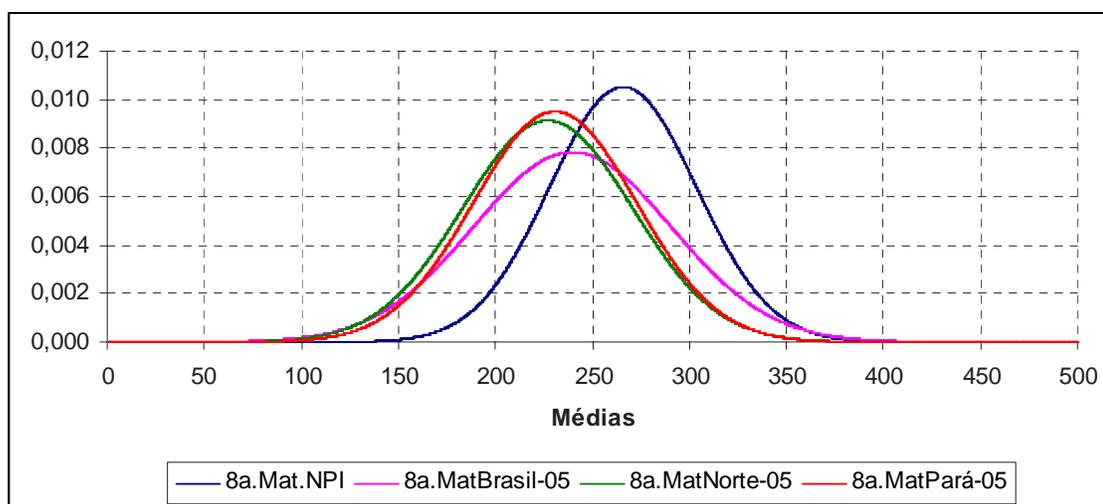


Figura 7.6 *Proficiências Médias em Matemática - 8ª série do Ensino Fundamental do NPI, do Brasil, da Região Norte e do Estado do Pará.*

### 7.2.2 Desempenho dos Alunos por Estágios de Competências

Conhecidas as proficiências dos alunos do NPI, foi possível calcular o percentual de estudantes em cada um dos níveis de desempenho do Saeb. Esses níveis são: *muito crítico*, *crítico*, *Intermediário* e *Adequado*.

De uma forma geral, 38% das crianças da 4ª série estão concentradas nos estágios *muito crítico* e *crítico* em Leitura, enquanto que em Matemática são 20% dos alunos que se encontram nesses estágios.

Para um aluno da 4ª série, estar no estágio *muito crítico* em Língua Portuguesa, significa não ter ainda desenvolvido as habilidades de leitura mínimas condizentes com quatro anos de escolarização. Quanto a prova de Matemática, esses alunos não são capazes de transpor para uma linguagem matemática específica, comandos operacionais elementares compatíveis com a série. O estudo mostrou que 21% dos alunos da 8ª série encontram-se nos estágios *muito crítico* e *crítico* em Leitura, enquanto que em Matemática, 30% dos alunos encontram-se nesses estágios.

A Figura 7.7 mostra o percentual dos alunos do NPI em cada um dos estágios, em relação às disciplinas : Matemática e Língua Portuguesa.

NPI - Matemática, 4ª Série			NPI - Língua Portuguesa, 4ª Série		
Níveis	Estágio	%	Níveis	Estágio	%
abaixo de 125	Muito Crítico	0,75	abaixo de 125	Muito Crítico	5,51
1 (125-150)	Crítico	3,76	1 (125-150)	Crítico	12,6
2 (150-175)	Crítico	15,8	2 (150-175)	Crítico	19,68
3 (175-200)	Intermediário	28,57	3 (175-200)	Intermediário	23,62
4 (200-250)	Intermediário	45,11	4 (200-250)	Intermediário	31,5
5 (250-300)	Adequado	4,51	5 (250-300)	Adequado	7,09
6 (300-350)	Adequado	1,5	6 (300-350)	Adequado	
	Total	100		Total	100

NPI - Matemática, 8ª Série			NPI - Língua Portuguesa, 8ª Série		
Níveis	Estágio	%	Níveis	Estágio	%
abaixo de 125	Muito Crítico	0	abaixo de 125	Muito Crítico	0,68
1 (125-150)	Muito Crítico	0	1 (125-150)	Muito Crítico	5,41
2 (150-175)	Muito Crítico	2,68	2 (150-175)	Crítico	8,78
3 (175-200)	Crítico	2,68	3 (175-200)	Crítico	6,08
4 (200-250)	Crítico	24,16	4 (200-250)	Intermediário	27,7
5 (250-300)	Intermediário	50,34	5 (250-300)	Intermediário	42,57
6 (300-350)	Intermediário	20,14	6 (300-350)	Intermediário	8,78
7 (350-375)	Adequado	0	7 (350-375)	Adequado	0
8 (375-400)	Adequado	0	8(375-mais)	Adequado	0
	Total	100		Total	100

Figura 7.7 Percentual de alunos por Estágios de Competências - NPI 2006

Nos Apêndices: C, D, E e F são mostradas as porcentagens de alunos do NPI, que dominam os diferentes níveis da escala do Saeb, tanto em Matemática como em Língua Portuguesa.

### 7.3 Conclusões

Segundo Araújo e Luzio (2005), a média mínima satisfatória para a 4.ª e a 8ª séries do Ensino Fundamental, em Língua Portuguesa são 200 e 300 pontos de proficiências respectivamente. O resultado do NPI em Língua Portuguesa para a 4.ª série foi 189,6 pontos e para a 8.ª série foi 242,3. Esses resultados estão abaixo do mínimo satisfatório, tornando-se evidente a necessidade de melhorar a qualidade dos processos de ensino-aprendizagem nas séries iniciais, garantindo os pré-requisitos necessários a um bom letramento e à com-

petência de leitura. Também para Matemática, a média mínima satisfatória para a 4.<sup>a</sup> e a 8.<sup>a</sup> séries do Ensino Fundamental são respectivamente 200 e 300 pontos de proficiência respectivamente. O resultado do NPI em Matemática para 4.<sup>a</sup> série foi 201,9 e para 8.<sup>a</sup> série foi 265,5. Pode-se dizer que o resultado dos alunos da 4.<sup>a</sup> série do Ensino Fundamental em Matemática está no mínimo satisfatório. Nesse Nível, as competências e as habilidades desenvolvidas pelos alunos da 4.<sup>a</sup> série irão permitir o progresso escolar com qualidade. Quanto aos alunos da 8.<sup>a</sup> série, o estudo mostra que não alcançaram o mínimo satisfatório.

## CONSIDERAÇÕES FINAIS

---

Neste trabalho apresentou-se uma visão geral sobre as principais avaliações desenvolvidas pelo Inep/MEC. Também apresentou-se de forma sucinta a principal ferramenta de avaliação educacional, a Teoria da Resposta ao Item, mas que também agrega ou pode ser usada em conjunto com outras metodologias, tais como a Teoria Clássica de Medidas (TCM), a Análise Fatorial (AF), a Análise Diferencial de Itens (DIF), a Análise de Desgaste de Itens (DRIFT), a Análise Hierárquica (HLM), dentre outras. Apresentou-se o processo de estimação por Máxima Verossimilhança Marginal, com abordagem de aplicação do algoritmo EM. Relativamente à estimação das proficiências, apresentou-se o procedimento de estimação pela Esperança da posteriori (EAP). Também apresentou-se o conjunto de níveis de proficiências adotados pelo MEC, referenciados como *muito crítico*, *crítico*, *Intermediário* e *Adequado*, ressaltando qual o domínio em Língua Portuguesa ou Matemática em cada faixa da escala de proficiência.

Em termos da aplicação com fins de colocar os resultados do NPI na escala do Saeb, tanto em Língua Portuguesa quanto em Matemática, aplicou-se um determinado caderno de prova fornecido pelo Inep, para cada série e disciplina, totalizado quatro cadernos. Também foram fornecidos os parâmetros de cada item, de forma que as proficiências foram estimadas considerando fixados (conhecidos) os parâmetros dos itens na escala (0,1). Posteriormente, as proficiências foram transformadas para a escala (250,50), em que são apresentados os resultados do Saeb.

Verificou-se que, no geral, o NPI está bem acima do desempenho nacional e do Estado do Pará, tanto em Língua Portuguesa quanto em Matemática. Como informação básica foram fornecidos os percentuais de alunos em cada nível da escala de proficiência. Por exemplo, enquanto no Brasil, segundo o Saeb-2005, a proporção de alunos da 4ª série em Matemática no nível *muito crítico* ou *crítico* é de 50%, no NPI este número é de 20,31%. Com estas informações em mãos, e do que é dominado em cada nível de proficiência, os professores

devem buscar estratégias que facilitem a compreensão dos alunos, incentivando-os ao interesse pela leitura, escrita e resolução de problemas do cotidiano, fazendo-os refletir, analisar, pensar e desenvolver a capacidade de raciocínio.

---

## Referências Bibliográficas

---

- [1] ANASTASI, A. *Psychological testing*. 6th ed. New York: Macmillan Publ. Co, 1988.
- [2] ANDERSEN, E. B. Conditional inference in multiple choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 1973.
- [3] ANDERSEN, E. B. *Discrete Statistical Models with Social Science Applications*. New York: North-Holland Publishing company, 1980.
- [4] ANDRADE, D. F. *Comparando o Desempenho de Grupos (Populações) de Respondentes Através da Teoria da Resposta ao Item*. Tese apresentada ao Departamento de Estatística e Matemática Aplicada da UFC para o concurso de professor titular, 1999.
- [5] ANDRADE, D. F. e Klein, R. Métodos Estatísticos para Avaliação Educacional: teoria da resposta ao item. *Boletim da ABE*, 43, 1999.
- [6] ANDRADE, D. F. e Valle, R. C. Introdução à teoria da resposta ao item: conceitos e aplicações. *Estudos em Avaliação Educacional*, 1998.
- [7] ANDRADE, D. F. TAVARES, H. R. e VALLE, R. C. *Teoria da Resposta ao Item: Conceitos e Aplicações*. ABE 2000.
- [8] ANDRICH, D. A rating formulation for ordered response categories. *Psychometrika*, 43, 1978.
- [9] ARAÚJO, C. H., LUZIO, N. Leitura na Educação básica. Brasília: Ministério da Educação, 25 jun. 2004. Disponível em <http://www.mec.gov.br/news/Artigos>, Acesso em 4 ago., 2004.
- [10] AZEVEDO, C. L. N., Método de Estimação na Teoria da Resposta ao Item. Dissertação de Mestrado. USP, São Paulo, 2003.
- [11] BAKER, F. B. *Item Response Theory - Parameter Estimation Techniques*. New York: Marcel Dekker, Inc., 1992.
- [12] BATISTA, J. R. *Valores Plausíveis para Estimação de Parâmetros Populacionais em Modelos da Teoria da Resposta ao Item*. Dissertação de Mestrado. Belo Horizonte: ICEX/UFMG, 1999.
- [13] BEATON, A. E. and ALLEN, N. L. Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17, 191-204, 1992.
- [14] BIRNBAUM, A. *Efficient design and use of tests of a mental ability for various decision-making problems*, ( Series Report No. 58-16. Project No. 7755-23). USAF School of Aviation Medicine, Texas: Randolph Air Force Base, 1957.

- [15] BIRNBAUM, A. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord & M.R. Novick. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley, 1968.
- [16] BLOOM, B. S., HASTINGS, J. T., MADAUS, G. F. *Evaluación del aprendizaje*. Buenos Aires: Troquel, 1975.
- [17] BLOOM, Benjamin, HASTINGS, Thomas e MADAUS, George. *Manual de avaliação formativa e somativa do aprendizado escolar*. São Paulo: Pioneira, 1983.
- [18] BOCK, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29-51, 1972.
- [19] BOCK, R. D. and LIEBERMAN, M. . Fitting a response model for  $n$  dichotomously scored items. *Psychometrika*, 35, 179-197, 1970.
- [20] BOCK, R. D. and AITKIN, M. *Marginal maximum likelihood estimation of item parameters: An application of a EM algorithm.*, 1981.
- [21] BOCK, R. D. and ZIMOWSKI, M. F. . Multiple Group IRT. In *Handbook of Modern Item Response Theory*. W. J. van der Linder e R.K. Hambleton Eds. New York: Springer-Verlag, 1997.
- [22] BRASIL. Lei de diretrizes e bases da educação nacional. Brasília: MEC/CNE, 1996.
- [23] CHOW, Y. S. and Teicher, H. *Probability Theory: Independence, Interchangeability, Martingales*. New York: Springer-Verlag, 1978.
- [24] DEMPSTER, A. P. , LAIRD, N. M. and RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, 39, 1-38, 1977.
- [25] DIAS SOBRINHO, José. *Universidade e Avaliação: entre a ética e o mercado*. Florianópolis: Insular, 2002.
- [26] ERTHAL, T. C. *Manual de Psicometria*. 7.ed. Rio de Janeiro: Jorge Zahar, 1955.
- [27] FLETCHER, P. R. *Teoria da Resposta ao Item: medidas invariantes do desempenho escolar. Ensaio: Avaliação e Políticas Públicas em Educação*. Rio de Janeiro, v. 1, p. 21-27, jan./mar. 1994.
- [28] FRANCISCO, R. *Aplicação da Teoria da Resposta ao Item (T.R.I.) no Exame Nacional de Cursos (E.N.C.) da Unicentro*. Dissertação (Mestrado em Ciências)- Universidade Federal do Paraná, Curitiba, 2005.
- [29] GOLDSTEIN, H. and Wood, R. Five decades of item response modelling. *British Journal of Mathematical and Statistical Psychology*, 42, 1989.
- [30] GRANGER, C. V., Deutsch, A. and LINN, R. T. Rasch analysis of the functional independence measure (FIMTM) mastery test. *Arch. Phys. Med. Rehabil.*, 1998.
- [31] GULLIKSEN, H. *Theory of Mental Tests*. New York : John Wiley and Sons, 1950.
- [32] HABERMAN, S. *Maximum Likelihood Estimates in Exponential Response Models*, (Technical Report) Chicago, IL: University of Chicago, 1975.

- [33] HAMBLTON, R. K. and SWAMINATHAN, H. *Item Response Theory: Principles and Applications*. Boston: Kluwer Academic Publishers, 1985.
- [34] HAMBLETON, R. K., SWAMINATHAN, H. and ROGERS, H. J. *Fundamentals of Item Response Theory*. Newbury Park : Sage Publications, 1991.
- [35] HAMBLETON, R. K. and COOK, L. L. Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, 14, 1997.
- [36] HOLANDA FERREIRA, A. *Novo Dicionário da Língua Portuguesa*. 1986
- [37] Instituto Nacional de Estudos e Pesquisas Educacionais Anízio Teixeira - INEP - disponível em <[www.inep.gov.br/download/catalogo\\_dinamico/enem/2001/documento\\_basico\\_enem\\_2001.pdf](http://www.inep.gov.br/download/catalogo_dinamico/enem/2001/documento_basico_enem_2001.pdf)>
- [38] Instituto Nacional de Estudos e Pesquisas Educacionais Anízio Teixeira - INEP - disponível em <[www.enceja.gov.br/](http://www.enceja.gov.br/)>
- [39] Instituto Nacional de Estudos e Pesquisas Educacionais Anízio Teixeira - INEP - disponível em <[www.inep.gov.br/internacional/pisa](http://www.inep.gov.br/internacional/pisa)>
- [40] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS *Relatório Saeb 2001 - Língua Portuguesa*. Brasília, 2002.
- [41] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS *Relatório Saeb 2001 - Matemática*. Brasília, 2002.
- [42] ISSAC, E. and KELLER, H. B. *Analysis of Numerical Methods*. New York: Wiley & Sons, 1966
- [43] KLEIN, R.; FONTANIVE, N. S. *Avaliação em larga escala: uma proposta inovadora*. Brasília, n. 66, p. 29-35, abr./jun., 1995.
- [44] LINDEN, W. J. van der and Hambleton, R. K. *Handbook of Modern Item Response Theory*. New York : Springer-Verlag, 1997.
- [45] LORD, F. M. A theory of test scores (No. 7). *Psychometric Monograph*, 1952.
- [46] LORD, F. M. An analysis of the verbal scholastic aptitude test using Birnbaum's three-parameter logistic model. *Educational and Psychological Measurement*, 28, 1968.
- [47] LORD, F. M. and NOVICK, M. R. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley, 1968.
- [48] LORD, F. M. Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, 39, 1974.
- [49] LORD, F. M. *Evaluation with artificial data of a procedure for estimating ability and item characteristic curve parameters*, (Research Bulletin RB-75-33). Princeton, NJ: Educational Testing Service, 1975.
- [50] LORD, F. M. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale: Lawrence Erlbaum Associates, Inc., 1980.
- [51] LUCKESI, C. C. *Avaliação de Avaliação Escolar*. 4 ed. São Paulo: Cortez, 1996.

- [52] LUCKESI, C. C. *Avaliação da aprendizagem escolar*. 17 ed. São Paulo: Cortez, 2005.
- [53] MARCOULIDES, G. A. Ed. *Modern Methods for Business Research*. Mahwah, NJ: Lawrence Erlbaum Reading, MA: Addison-Wesley, 1998.
- [54] MASTERS, G. N. A Rasch model for partial credit scoring. *Psychometrika*, 47, 1982.
- [55] Ministério da Educação e do Desporto. *Sistema Nacional de Avaliação da Educação Básica: SAEB 2003 - relatório: Análise de Teoria da Resposta ao Item*. Rio de Janeiro: Fundação Cesgranrio, 2004.
- [56] Ministério da Educação e do Desporto. *Sistema Nacional de Avaliação da Educação Básica: SAEB 2003 - relatório: Análise Clássica dos Testes* Rio de Janeiro : Fundação Cesgranrio, 2004.
- [57] MISLEVY, R. J. Bayes modal estimation in item response models. *Psychometrika*, 51, 177-195, 1986a.
- [58] MISLEVY, R. J. Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196, 1991.
- [59] MISLEVY, R. J. *Linking Educational Assessments : concepts, issues, methods and prospects*. Princeton : Educational Testing Service, 1992.
- [60] MISLEVY, R. J. and STOCKING, M. L. A Consumer's Guide to LOGISTIC and BILOG. *Applied Psychological Measurement*, 13 57-75, 1989.
- [61] MISLEVY, R. J. and BOCK, R. D. *BILOG 3 : Item Analysis and Test Scoring with Binary Logistic Models*. Chicago : Scientific Software, Inc. 1990.
- [62] MURAKI, E. A generalized partial credit model : Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159-176, 1992.
- [63] MURAKI, E. and BOCK, R. D. *PARSCALE : IRT Based Test Scoring and Item Analysis for Graded Open-Ended Exercises and Performance Tasks*. Chicago: Scientific Software, Inc., 1997.
- [64] NOVICK, M. R. The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3,1-18.1966
- [65] Organization for Economic Co-operation and Development(OECD). *Knowledge and Skills for Life: First results from PISA 2000*, 2001.
- [66] PASQUALI, L. *Psicometria: Teoria dos testes na Psicologia e na Educação*,Ed. Vozes, Petrópolis, 2003.
- [67] RASCH, G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen : Danish Institute for Educational Research, 1960.
- [68] RICHARDSON, M. W. The relationship between difficulty and the differential validity of a test. *Psychometrika*, 1, 33-49, 1936.
- [69] SAMEJIMA, F. A. Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17.,1969.

- [70] *Sistema Nacional de Avaliação da Educação Básica - SAEB* - disponível em <[www.inep.gov.br/download/saeb/2001/relatorioSAEB\\_matematica.pdf](http://www.inep.gov.br/download/saeb/2001/relatorioSAEB_matematica.pdf)>
- [71] *Sistema Nacional de Avaliação da Educação Básica - SAEB* - disponível em <[www.inep.gov.br/download/saeb/2001/relatorioSAEB\\_portugues.pdf](http://www.inep.gov.br/download/saeb/2001/relatorioSAEB_portugues.pdf)>
- [72] SOEIRO, Leda e AVELINE, Suely. Avaliação Educacional. *Avaliação Educacional*. Porto Alegre: Sulina, 1982.
- [73] STROUD, A. H. and SECREST, D. *Gaussian Quadrature Formulas*. Englewood Cliffs, New Jersey : Prentice-Hall, 1966.
- [74] URBINA, S. *Fundamentos da testagem Psicológicas*. Ed; Artmed, Porto Alegre, 2007.
- [75] VALLE, R. C. *Teoria da Resposta ao Item..* Dissertação de Mestrado. São Paulo: IME/USP, 1999.
- [76] ZIMOWSKI, M. F. MURAKI, E. , MISLEVY, R. J. and Bock, R.D. *BILOG-MG: Multiple-Group IRT Analysis and Test Maintenance for Binary Items*. Chicago: Scientific Software, Inc., 1996.

---

## Apêndice A

# Matriz de Referência para Língua Portuguesa e Matemática

---

Matriz de Referência para Língua Portuguesa:

- Procedimentos de leitura
  - D1 - Localizar informações explícitas em um texto.
  - D3 - Inferir o sentido de uma palavra ou expressão.
  - D4 - Inferir uma informação implícita em um texto.
  - D6 - Identificar o tema de um texto.
  - D11 - Distinguir um fato da opinião relativa a esse fato.
- Implicações do suporte, do gênero e/ou do enunciador na compreensão do texto
  - D5 - Interpretar texto com auxílio de material gráfico diverso (propagandas, quadros, fotos, etc.),
  - D9 - Identificar a finalidade de textos de diferentes gêneros.
- Relação entre textos
  - D15 - Reconhecer diferentes formas de tratar uma informação na comparação de textos que tratam do mesmo tema, em função das condições em que ele foi produzido e daquelas em que será recebido.
- Coerência e coesão no processamento do texto. D2 - Estabelecer relações entre partes de um texto, identificando repetições ou substituições que contribuem para a continuidade de um texto,
  - D7 - Identificar o conflito gerador do enredo e os elementos que constroem a narrativa.
  - D8 - Estabelecer relação causa e consequência entre partes e elementos do texto.
  - D12 - Estabelecer relações lógico-discursivas presentes no texto, marcadas por conjunções, advérbios. etc.
- Relações entre recursos expressivos e efeitos de sentido;
  - D13 - Identificar efeitos de ironia ou humor em textos variados.
  - D14 - Identificar o efeito de sentido decorrente do uso da pontuação e de outras notações.
- Variação lingüística.
  - D10 - Identificar as marcas lingüísticas que evidenciam o locutor e o interlocutor de um texto.

Matriz de Referência para Matemática.

- Espaço e Forma.
  - D1 - Identificar a localização /movimentação de objeto em mapas, croquis e outras representações gráficas.
  - D2 - Identificar propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações.
  - D3 - Identificar propriedades comuns e diferenças entre figuras bidimensionais pelo

número de lados, pelos tipos de ângulos.

D4 - Identificar quadriláteros observando as posições relativas entre seus lados (paralelos, concorrentes, perpendiculares).

D5 - Reconhecer a conservação ou modificação de medidas dos lados, do perímetro, da área em ampliação e /ou redução de figuras poligonais usando malhas quadriculadas.

- Grandezas e Medida.

D6 - Estimar a medida de grandezas utilizando unidades de medida convencionais ou não.

D7 - Resolver problemas significativos utilizando unidades de medida padronizadas como km/m/cm/mm, kg/g/mg, l/ml.

D8 - Estabelecer relações entre unidades de medida de tempo.

D9 - Estabelecer relações entre o horário de início e término e /ou o intervalo da duração de um evento ou acontecimento.

D10 - Num problema, estabelecer trocas entre cédulas e moedas do sistema monetário brasileiro, em função de seus valores.

D11 - Resolver problema envolvendo o cálculo do perímetro de figuras planas, desenhadas em malhas quadriculadas.

D12 - Resolver problema envolvendo o cálculo ou estimativa de áreas de figuras planas, desenhadas em malhas quadriculadas.

- Números e Operações /Álgebra e Funções.

D13 - Reconhecer e utilizar características do sistema de numeração decimal, tais como agrupamentos e trocas na base 10 e princípio do valor posicional.

D14 - Identificar a localização de números naturais na reta numérica.

D15 - Reconhecer a decomposição de números naturais nas suas diversas ordens.

D16 - Reconhecer a composição e a decomposição de números naturais em sua forma polinomial.

D17 - Calcular o resultado de uma adição ou subtração de números naturais.

D18 - Calcular o resultado de uma multiplicação ou divisão de números naturais.

D19 - Resolver problema com números naturais, envolvendo diferentes significados da adição ou subtração: juntar, alteração de um estado inicial (positiva ou negativa), comparação e mais de uma transformação (positiva ou negativa).

D20 - Resolver problema com números naturais, envolvendo diferentes significados da multiplicação ou divisão: multiplicação comparativa, idéia de proporcionalidade, configuração retangular e combinatória.

D21 - Identificar diferentes representações de um mesmo número racional.

D22 - Identificar a localização de números racionais representados na forma decimal na reta numérica.

D23 - Resolver problema utilizando a escrita decimal de cédulas e moedas do sistema monetário brasileiro.

D24 - Identificar fração como representação que pode estar associada a diferentes significados.

D25 - Resolver problema com números racionais expressos na forma decimal envolvendo diferentes significados da adição ou subtração.

D26 - Resolver problema envolvendo noções de porcentagem (25

- Tratamento da Informação.

D27 - Ler informações e dados apresentados em tabelas.

D28 - Ler informações e dados apresentados em gráficos (particularmente em gráficos de colunas).

Apêndice B

**Brasil - Proficiências do Saeb 1995 - 2005**

Tabela B.1 *Média de Proficiências em Língua Portuguesa*

<i>Série</i>	<i>1995</i>	<i>1997</i>	<i>1999</i>	<i>2001</i>	<i>2003</i>	<i>2005</i>	<i>Dif.</i>	<i>Sig.</i>
4a.Série do E.F.	188,3 (1,6) <sup>e</sup>	186,5 (1,6)	170,7 (0,9)	165,1 (0,8)	169,4 (0,8)	172,3 (1,0)	2,9	*
8a.Série do E.F.	256,1 (1,4)	250,0 (2,0)	232,9 (1,0)	235,2 (1,3)	232,0 (1,0)	231,9 (1,0)	-0,1	
3a.Série do E.M.	290,0 (1,9)	283,9 (2,1)	266,6 (1,5)	262,3 (1,4)	266,7 (1,3)	257,6 (1,6)	-9,1	*

*e*: erro padrão

Fonte: Saeb, 2005

Tabela B.2 *Média de Proficiências em Matemática*

<i>Série</i>	<i>1995</i>	<i>1997</i>	<i>1999</i>	<i>2001</i>	<i>2003</i>	<i>2005</i>	<i>Dif.</i>	<i>Sig.</i>
4a.Série do E.F.	190,6 (1,5) <sup>e</sup>	190,8 (1,2)	181,0 (0,9)	176,3 (0,8)	171,1 (0,8)	182,4 (1,9)	5,3	*
8a.Série do E.F.	253,2 (2,9)	250,0 (2,1)	246,4 (1,1)	243,4 (1,2)	245,0 (1,1)	239,5 (1,1)	-5,5	*
3a.Série do E.M.	281,9 (2,6)	288,7 (3,0)	280,3 (1,7)	276,7 (1,3)	278,7 (1,4)	271,3 (1,8)	7,4	*

*e*: erro padrão

Fonte: Saeb, 2005

Obs: Tanto para os resultados de Língua Portuguesa como os de Matemática, as médias dos anos de 1995, 2003 e 2005 foram estimadas incluindo o estrato de escolas públicas federais. Em todos os anos, a zona rural foi avaliada e incluída para a estimativa das médias apenas na 4ª série. Para a composição do estrato rural não foi incluída a Região Norte em 1997 e em 1999 e 2001, apenas participaram os estados da Região Nordeste, Minas Gerais e o Mato Grosso.

Apêndice C

## Escala de Desempenho - Matemática, 4ª série

(Adaptado do Relatório Saeb 2001- Matemática)

ESCALA DE DESEMPENHO – Matemática, 4ª Série  
 Na 4ª série do Ensino Fundamental, os alunos dominam as seguintes habilidades:

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		1 3,76% 125 - 150	2 15,78% 150-175	3 28,57% 175-200	4 45,11% 200 - 250	5 4,51% 250 - 300	6 1,51% 300 - 350
<b>Espaço e Forma</b>	Calculam áreas	• de figuras desenhadas em malhas quadriculadas					
	Reconhecem a localização ou movimentação de objetos.	<ul style="list-style-type: none"> <li>• situado em seu referencial, em representações gráficas.</li> <li>• em referencial diferente do seu, em representações gráficas.</li> </ul>					
	Distinguem sólidos	<ul style="list-style-type: none"> <li>• que têm superfície arredondada.</li> <li>• com forma esférica quadrado fora da posição usual e elementos de figuras tridimensionais (face, vértice, aresta)</li> </ul>					
	Utilizam propriedades dos polígonos.	•identificando seus lados					

(continuação)

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		1 3,76% 125 - 150	2 15,78% 150-175	3 28,57% 175-200	4 45,11% 200 - 250	5 4,51% 250 - 300	6 1,51% 300 - 350
Grandezas e Medidas	Estimam valores	<ul style="list-style-type: none"> <li>de uma mesma medida</li> </ul>					
	Lêem horas e minutos	<ul style="list-style-type: none"> <li>em relógio digital               <ul style="list-style-type: none"> <li>em relógios de ponteiros                   <ul style="list-style-type: none"> <li>convertendo a duração de tempo.</li> </ul> </li> </ul> </li> </ul>					
	Resolvem problemas	<ul style="list-style-type: none"> <li>envolvendo intervalos de tempo em situações cotidianas.               <ul style="list-style-type: none"> <li>com intervalo de tempo (meses)</li> </ul> </li> </ul>					
	Identificam moedas	<ul style="list-style-type: none"> <li>para trocar uma quantia pequena em dinheiro.</li> </ul>					
	Realizam conversão	<ul style="list-style-type: none"> <li>entre medidas de tempo (hora, dia, semana)               <ul style="list-style-type: none"> <li>de kg para g; de l para ml                   <ul style="list-style-type: none"> <li>de m para cm.</li> </ul> </li> </ul> </li> </ul>					
	Calculam perímetro	<ul style="list-style-type: none"> <li>de figura poligonal conhecendo a medida dos lados.</li> </ul>					
	Reconhecem unidade de medida de área	<ul style="list-style-type: none"> <li>como o metro quadrado</li> </ul>					
Estimam comprimento	<ul style="list-style-type: none"> <li>usando medidas não convencionais.</li> </ul>						

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		1 3,76% 125 - 150	2 15,80% 150-175	3 28,57% 175-200	4 45,11% 200 - 250	5 4,51% 250 - 300	6 1,51% 300 - 350
Números e Operações	Reconhecem Partes de um todo	<ul style="list-style-type: none"> <li>em representações gráficas</li> </ul>					
	Resolvem problemas	<ul style="list-style-type: none"> <li>do cotidiano envolvendo adições de pequenas quantias de dinheiro.</li> <li>do cotidiano envolvendo adição de números naturais e racionais com o mesmo número de casas decimais.               <ul style="list-style-type: none"> <li>envolvendo subtrações de números racionais com o mesmo número de casas</li> <li>simples com as quatro operações.                   <ul style="list-style-type: none"> <li>que utilizam a multiplicação envolvendo a noção de proporcionalidade.</li> </ul> </li> <li>envolvendo mais de uma operação, incluindo o sistema monetário.</li> <li>de composição e decomposição mais complexos (mais zeros e ordens).                   <ul style="list-style-type: none"> <li>envolvendo divisão com resto ou multiplicação combinatória</li> <li>envolvendo proporcionalidade</li> </ul> </li> </ul> </li> </ul>					
	Calculam o resultado	<ul style="list-style-type: none"> <li>de uma adição e de uma subtração envolvendo números de até três algarismos, inclusive com recuo e reserva.</li> <li>de multiplicações com um algarismo.</li> </ul>					
	Identificam frações	<ul style="list-style-type: none"> <li>com apoio de representação gráfica.</li> </ul>					
	Reconhecem o valor posicional	<ul style="list-style-type: none"> <li>do Sistema de Numeração Decimal</li> </ul>					
	Decompõem um número natural	<ul style="list-style-type: none"> <li>em suas ordens e vice-versa.               <ul style="list-style-type: none"> <li>em agrupamento de 1000.</li> </ul> </li> </ul>					
	Calculam o resultado	<ul style="list-style-type: none"> <li>de subtrações mais complexas com números naturais.</li> <li>de multiplicações com números de 2 algarismos.</li> <li>de divisões exatas por números de 1 algarismo               <ul style="list-style-type: none"> <li>de uma divisão por números</li> <li>de 2 algarismos, inclusive com restos.</li> <li>de porcentagens simples.</li> </ul> </li> </ul>					
	Identificam número natural	<ul style="list-style-type: none"> <li>em intervalo da reta numérica.</li> </ul>					
	Reconhecem a composição/decomposição	<ul style="list-style-type: none"> <li>na escrita decimal, em casos mais complexos.</li> </ul>					
	Comparam números racionais	<ul style="list-style-type: none"> <li>na forma decimal.</li> </ul>					

(continuação)

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		1 3,76% 125 - 150	2 15,80% 150 - 175	3 28,57% 175 - 200	4 45,11% 200 - 250	5 4,51% 250 - 300	6 1,51% 300 - 350
Tratamento da Informação	Lêem informações	<ul style="list-style-type: none"> <li>em tabelas e gráficos de colunas</li> </ul>					
	Interpretam gráfico	<ul style="list-style-type: none"> <li>de colunas através de leitura de valores no eixo vertical               <ul style="list-style-type: none"> <li>de seletores, associando a dados de uma tabela.</li> <li>de linhas correspondentes a uma seqüência de valores (posit. e negativos)</li> </ul> </li> </ul>					
	Resolvem problemas	<ul style="list-style-type: none"> <li>simples, usando dados em gráficos de barras ou tabelas.</li> </ul>					

Apêndice D

## Escala de Desempenho - Matemática, 8ª série

ESCALA DE DESEMPENHO – MATEMÁTICA, 8ª Série

Na 8ª Série do Ensino Fundamental, além das habilidades descritas para a 4ª série, os alunos também dominam as seguintes habilidades:

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS				
		4 24,16% 200-250	5 50,34% 250-300	6 20,14% 300-350	7 0,0% 350-375	8 0,0% 375-400
Espaço e Forma	Localizam pontos	<ul style="list-style-type: none"> <li>• usando coordenadas em um referencial quadriculado.                             <ul style="list-style-type: none"> <li>• no plano cartesiano.</li> </ul> </li> </ul>				
	Identificam planificações	<ul style="list-style-type: none"> <li>• de um sólido simples dado em perspectiva de um cilindro dado em situação concreta (lata de óleo).</li> </ul>				
	Estabelecem diferenças	<ul style="list-style-type: none"> <li>• entre ângulos, quadrados e círculos.</li> </ul>				
	Identificam lados e ângulos	<ul style="list-style-type: none"> <li>• de um quadrilátero (retângulo, losango, quadrado e trapézio).</li> </ul>				
	Reconhecem medidas do perímetro	<ul style="list-style-type: none"> <li>• de um retângulo em malha quadriculada.</li> </ul>				
	Calculam volumes	<ul style="list-style-type: none"> <li>• através de contagem de blocos.</li> </ul>				
	Avaliam distâncias	<ul style="list-style-type: none"> <li>• horizontais e verticais em croquis usando escalas gráficas.</li> </ul>				
	Classificam ângulos	<ul style="list-style-type: none"> <li>• em agudos, retos ou obtusos, de acordo com a sua medida em graus.</li> </ul>				
	Resolvem problema	<ul style="list-style-type: none"> <li>• envolvendo ângulos agudos, usando inclusive a lei angular de Thales e aplicando o Teorema de Pitágoras.                             <ul style="list-style-type: none"> <li>• utilizando a propriedade de semelhança de triângulos, de quadrilátero (como por exemplo, o Tangram).</li> </ul> </li> </ul>				
	Reconhecem a definição de circunferência	<ul style="list-style-type: none"> <li>• e seus elementos (raio, diâmetro e corda).</li> </ul>				
Calculam áreas	de figuras simples (triângulos, paralelogramos, retângulos e trapézios)					

continuação

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS				
		4 24,16% 200-250	5 50,34% 250-300	6 20,14% 300-350	7 0,0% 350-375	8 0,0% 375-400
<b>Grandezas e Medidas</b>	Operam com as unidades de medidas	<ul style="list-style-type: none"> <li>• não convencionais e reconhecem que quanto maior a medida, menor a unidade.</li> </ul>				
	Resolvem problemas de conversão de unidade de medida	<ul style="list-style-type: none"> <li>• como, por exemplo, metro cúbico em litro.</li> </ul>				
	Calculam área	<ul style="list-style-type: none"> <li>• de regiões poligonais desenhadas em malhas quadriculadas.</li> </ul>				
	Calculam volumes	<ul style="list-style-type: none"> <li>• de um bloco retangular</li> </ul>				
<b>Números e Operações</b>	Localizam na reta numérica	<ul style="list-style-type: none"> <li>• números inteiros, positivos, negativos e números racionais na forma decimal.</li> </ul>				
	Identificam o sistema de equações	<ul style="list-style-type: none"> <li>• de primeiro grau, expressas em uma situação dada.</li> </ul>				
	Calculam resultados de operações	<ul style="list-style-type: none"> <li>• de adição com números racionais e com diferentes casas decimais.</li> <li>• de potenciação com números inteiros, positivos e negativos.</li> <li>• de transformação de fração em porcentagens e vice-versa.</li> </ul>				
	Resolvem problemas	<ul style="list-style-type: none"> <li>• simples envolvendo frações e porcentagens.</li> <li>• de equação do segundo grau.</li> <li>• que envolvem o conceito de proporcionalidade.</li> <li>• de juros simples.</li> </ul>				
	Resolvem expressões	<ul style="list-style-type: none"> <li>• Com números inteiros positivos e negativos e também com números racionais.</li> <li>• envolvendo as quatro operações, potências e raízes.</li> <li>• Com números inteiros positivos e negativos sem que os sinais estejam explicitados.</li> </ul>				
	Comparam números racionais	<ul style="list-style-type: none"> <li>• usando arredondamento.</li> </ul>				
	Ordenam números	<ul style="list-style-type: none"> <li>• Inteiros, positivos e negativos e identificam o intervalo onde se encontra uma raiz quadrada não exata.</li> </ul>				

continuação

TEMA	HABILIDADES	NÍVEIS E PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS				
		4 24,16% 200-250	5 50,34% 250-300	6 20,14% 300-350	7 0,0% 350-375	8 0,0% 375-400
Tratamento da Informação	Lêem tabelas	<ul style="list-style-type: none"> <li>de dupla entrada, como, por exemplo, dados de peso e altura.               <ul style="list-style-type: none"> <li>com números positivos e negativos e identificam o gráfico de colunas correspondentes.</li> </ul> </li> </ul>				
	Reconhecem gráficos	<ul style="list-style-type: none"> <li>de colunas referentes a dados apresentados de forma textual.               <ul style="list-style-type: none"> <li>envolvendo regiões do plano cartesiano.</li> <li>de colunas relativos a um gráfico de setor.</li> <li>de linhas com duas seqüências de valores.</li> <li>e estimam quantidades em gráficos de setores.</li> </ul> </li> </ul>				
	Resolvem problemas	<ul style="list-style-type: none"> <li>de comparação entre gráficos de coluna.</li> <li>de cálculo da média aritmética de um conjunto de valores.</li> </ul>				

Apêndice E

## Escala de Desempenho - Língua Portuguesa, 4ª série

ESCALA DE DESEMPENHO – LÍNGUA PORTUGUESA, 4ª Série  
Na 4ª Série do Ensino Fundamental os alunos dominam as seguintes habilidades:

TÓPICOS	HABILIDADES	NÍVEIS PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS				
		1 12,6% 125-150	2 19,68% 150-175	3 23,62% 175-200	4 31,5% 200-250	5 7,09% 250-300
Procedimentos de Leitura	Localizam informações explícitas	<ul style="list-style-type: none"> <li>• em um texto</li> <li>• em textos narrativos curtos como histórias infantis (conto de fada e fábulas).</li> <li>• em outros gêneros textuais como comunicado da escola aos pais e pequenos poemas descritivos e narrativos.                             <ul style="list-style-type: none"> <li>• em textos mais complexos mais longos e não narrativos (textos publicitários).</li> <li>• em textos publicitários de revistas e jornais.</li> </ul> </li> <li>• em textos poéticos não descritivos (poemas simples e canção popular)</li> </ul>				
	Inferem o sentido	<ul style="list-style-type: none"> <li>• de palavras em textos simples (história em quadrinhos)                             <ul style="list-style-type: none"> <li>• de palavras ou expressão em texto narrativo (história infantil).</li> <li>• de expressão de uso corrente em textos informativos.</li> </ul> </li> </ul>				
	Identificam informações implícitas	<ul style="list-style-type: none"> <li>• acompanhada de ilustrações em histórias em quadrinhos e em textos narrativos simples (prosa ou em verso)                             <ul style="list-style-type: none"> <li>• em textos descritivos (fábula).</li> <li>• em narrativa curta com material ilustrativo.</li> </ul> </li> </ul>				
	Identificam o tema	<ul style="list-style-type: none"> <li>• de textos narrativos simples (histórias em infantis) com apoio de elementos (figuras)                             <ul style="list-style-type: none"> <li>• de um texto informativo (em revista).</li> <li>• de texto poético de baixa complexidade (poema descritivo).</li> <li>• em narrativa curta (fábula) com base em material ilustrativo.</li> </ul> </li> </ul>				
	Distinguem fato de opinião relativa a esse fato	<ul style="list-style-type: none"> <li>• em textos narrativos simples (histórias em quadrinhos) e textos com ou sem ilustrações.</li> </ul>				

TÓPICOS	HABILIDADES	NÍVEIS PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS				
		1 12,6% 125-150	2 19,68% 150-175	3 23,62% 175-200	4 31,5% 200-250	5 7,09% 250-300
<b>Implicações do Suporte, do Gênero e/ou do Enunciador na Compreensão do Texto</b>	Interpretam texto	<ul style="list-style-type: none"> <li>Com auxílio de recursos gráficos (em histórias em quadrinhos).</li> </ul>				
	Identificam a finalidade	<ul style="list-style-type: none"> <li>de textos narrativos simples.</li> <li>de textos de diferentes gêneros.</li> </ul>				
<b>Coerência e Coesão em Processamento do Texto</b>	Identificam elementos	<ul style="list-style-type: none"> <li>que constroem uma narrativa simples (espaço) <ul style="list-style-type: none"> <li>constitutivos da narrativa (personagens).</li> <li>que identificam o conflito gerador do enredo. <ul style="list-style-type: none"> <li>espaço (em crônicas), personagens, conflito e desfecho (em histórias infantis).</li> </ul> </li> </ul> </li> </ul>				
	Estabelecem relações de continuidade	<ul style="list-style-type: none"> <li>em histórias em quadrinhos</li> <li>por meio de relações anafóricas em textos curtos e simples. <ul style="list-style-type: none"> <li>Por meio de relações anafóricas e textos de complexidade média.</li> </ul> </li> </ul>				
	Estabelecem relações causa consequência	<ul style="list-style-type: none"> <li>entre partes e elementos de textos. <ul style="list-style-type: none"> <li>em poemas curtos. <ul style="list-style-type: none"> <li>em textos poéticos.</li> <li>em textos de complexidades média (fábula).</li> </ul> </li> </ul> </li> </ul>				
	Estabelecem relações lógico discursivas	<ul style="list-style-type: none"> <li>presentes no texto marcadas por conjunções.</li> </ul>				
<b>Relações entre Recursos Expressivos e Efeitos de Sentido</b>	Identificam o efeito de sentido	<ul style="list-style-type: none"> <li>decorrente da disposição gráfica das palavras em um texto. <ul style="list-style-type: none"> <li>decorrente do uso expressivo da pontuação (ponto de interrogação) e De repetição de uma letra em uma palavra.</li> </ul> </li> </ul>				
<b>Variação Lingüística</b>	Identificam marcas lingüísticas	<ul style="list-style-type: none"> <li>próprias do meio rural e urbano <ul style="list-style-type: none"> <li>que caracterizam o locutor e o interlocutor do texto.</li> </ul> </li> <li>próprias de textos comerciais.</li> </ul>				

Apêndice F

## Escala de Desempenho - Língua Portuguesa, 8ª série

ESCALA DE DESEMPENHO – LÍNGUA PORTUGUESA, 8ª Série  
 Na 8ª Série do Ensino Fundamental, além das habilidades descritas para a 4ª série, os alunos dominam as seguintes habilidades:

TÓPICOS	HABILIDADES	NÍVEIS PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		2 8,78% 150-175	3 6,08% 175-200	4 27,70% 200-250	5 42,57% 250-300	6 8,78% 300-350	7 0,0% 350-375
<b>Procedimentos de Leitura</b>	Localizam informações explícitas	<ul style="list-style-type: none"> <li>em textos mais complexos (notícias de revistas e jornais).</li> <li>comuns em dois textos de jornais.</li> </ul>					
	Inferem o sentido	*	*			<ul style="list-style-type: none"> <li>de informações contidas em gráficos e tabelas.</li> <li>em textos de baixa complexidade.</li> </ul>	
	Identificam informações implícitas	*	*	*		<ul style="list-style-type: none"> <li>que implica conceitos mais abstratos.</li> <li>em texto poético de maior complexidade.</li> </ul>	
	Identificam o tema	*	*		<ul style="list-style-type: none"> <li>em textos poéticos.</li> </ul>	<ul style="list-style-type: none"> <li>e os sentidos metafóricos em textos narrativos longos (contos) e em material prático.</li> </ul>	
	Distinguem fato de opinião relativa a esse fato	*	*	*	*	<ul style="list-style-type: none"> <li>em matéria jornalística (texto informativo contendo análise de dados)</li> </ul>	
<b>Implicações do Suporte, do Gênero e/ou do Enunciador na Compreensão do Texto</b>	Interpretam texto	*	*	*		<ul style="list-style-type: none"> <li>com auxílio de material gráfico diverso e em vários níveis de abstração que extrapolam o sentido convencional.</li> </ul>	
	Identificam a finalidade		<ul style="list-style-type: none"> <li>de um *</li> </ul>	<ul style="list-style-type: none"> <li>texto informativo *</li> </ul>			
<b>Redação entre Textos</b>	Reconhecem diferentes formas de tratar a informação sobre o mesmo tema				<ul style="list-style-type: none"> <li>como texto narrativo e provérbios.</li> </ul>	<ul style="list-style-type: none"> <li>como no caso de um boletim metodológico com informações pictóricas em tabelas e gráficos.</li> </ul>	

\* Habilidades descritas na escala de 4ª série do Ensino Fundamental e construída, também, pelos de 8ª série em cada nível.

TÓPICOS	HABILIDADES	NÍVEIS PERCENTUAIS DE ALUNOS DO NPI NOS NÍVEIS					
		2 8,78% 150-175	3 6,08% 175-200	4 27,70% 200-250	5 42,57% 250-300	6 8,78% 300-350	7 0,0% 350-375
<b>Coerência e Coesão no Procedimento do Texto</b>	Estabelecem relações de continuidade						<ul style="list-style-type: none"> <li>identificando repetições e substituição em um texto.</li> <li>relacionando uma palavra de sentido mais genérico a outra de sentido mais específico.               <ul style="list-style-type: none"> <li>reconhecendo a mesma idéia expressa em duas partes de um mesmo texto, com apoio de figuras.</li> <li>reconhecendo a transformação de referente ao longo do texto narrativo</li> </ul> </li> </ul>
	Estabelecem relações entre teses e argumentos	*	*				<ul style="list-style-type: none"> <li>em texto simples (bilhetes).</li> </ul>
	Identificam conflito gerador do enredo e elementos que constroem a narrativa	*	*	*			<ul style="list-style-type: none"> <li>distinguem narrador e personagem em pequenos textos.</li> </ul>
	Estabelecem relações causa/conseqüência	*	*				<ul style="list-style-type: none"> <li>em textos poéticos.</li> <li>com base em pistas léxicas</li> </ul>
	Estabelecem relações Lógico-discursivas	*	*	*	*		<ul style="list-style-type: none"> <li>mais complexas pelo uso de conjugações.</li> </ul>
<b>Relações entre Recursos Expressivos e Efeito de Sentido</b>	Identificam efeito de ironia ou humor em textos variados	*	*	*			<ul style="list-style-type: none"> <li>em história em quadrinhos pelo uso de recursos morfossintáticos (humor)</li> <li>em textos poéticos descritivos simples (ironia).</li> </ul>
	Identificam efeito de sentido						<ul style="list-style-type: none"> <li>gerado pela repetição de sons e palavras em textos poéticos descritivos.</li> <li>pela relação forma/conteúdo em textos poéticos de autores modernistas.</li> </ul>
	Identificam efeito de exploração de recursos						<ul style="list-style-type: none"> <li>visuais (tamanho e forma de letras)</li> <li>como o uso de caixa alta para enfatizar uma palavra.</li> </ul>
<b>Variação Lingüística</b>	Identificam marcas lingüística						<ul style="list-style-type: none"> <li>que evidenciam o locutor e o interlocutor em texto informativo.</li> <li>que caracterizam a linguagem formal e informal, e o nível de adequação do uso em relação ao interlocutor.</li> </ul>

## Método da Quadratura

---

Para calcular o valor aproximado da integral definida utiliza-se uma combinação linear de valores da função  $f$  em certos pontos  $x_i$ ,  $a \leq x_i \leq b$ , chamados de nós, e de certos valores  $w_i$ ,  $i = 1, 2, 3, \dots, n + 1$  que constituem os pesos, ou seja :

$$\int_a^b f(x)dx \approx w_1f(x_1) + w_2f(x_2) + w_3f(x_3) + \dots + w_{1n}f(x_n) + w_{n+1}f(x_{n+1}) \quad (\text{G.1})$$

A expressão (G.1) corresponde a aproximar a integral pela soma da áreas de retângulos de altura  $f(x_i)$  e largura  $w_i$ . De acordo com os valores dos pesos e com a escolha de nós, temos no lado direito de (G.1) o que se costuma chamar de regras de integração, quadraturas ou soma de quadratura.

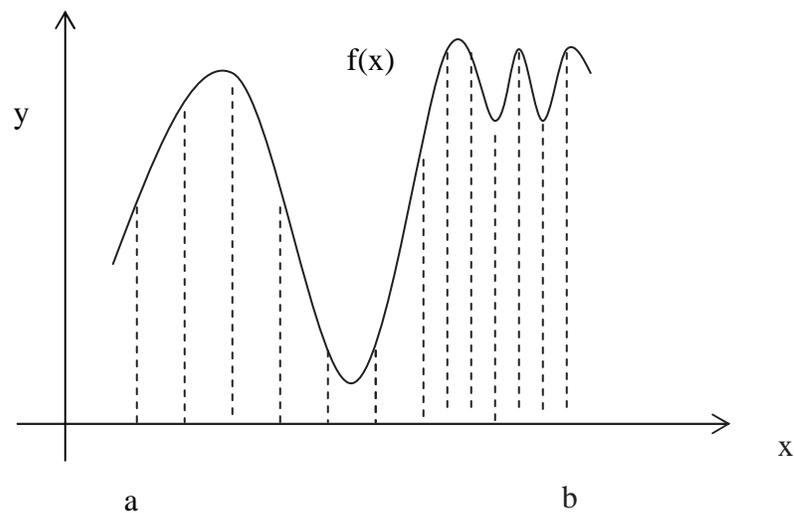


Figura G.1