

**ANÁLISE COMPARATIVA DOS PROCESSOS DE AVALIAÇÃO
EDUCACIONAL EM LARGA ESCALA**

Mauro Luiz Rabelo¹

¹ Departamento de Matemática, Universidade de Brasília, Brasil, *e-mail*: rabelo@unb.br

Índice

Introdução	1
1. Os Exames Nacionais de Avaliação Educacional	3
1.1. Contextualização: do SAEB à Prova Nacional de Concurso para o Ingresso na Carreira Docente	3
1.2. O Sistema de Avaliação da Educação Básica (SAEB) e a Prova Brasil	4
1.2.1. O que é avaliado nas provas	8
1.2.2. Matrizes de Referência de Matemática	10
1.2.3. Análise do desempenho dos estudantes – aplicação da TRI	14
1.2.4. Interpretação dos níveis da escala	19
1.2.5. Construção da série histórica de desempenho	23
1.2.6. O Índice de Desenvolvimento da Educação Básica (IDEB)	24
1.3. O Exame Nacional do Ensino Médio (ENEM)	26
1.3.1. Características do Exame	29
1.3.2. Matriz de Referência para o ENEM 2009 – Matemática e suas Tecnologias	31
1.3.3. Objetos de conhecimento de Matemática e suas Tecnologias	33
1.4. O Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA)	34
1.5. O Sistema Nacional de Avaliação da Educação Superior (SINAES)	35
1.5.1. O Exame Nacional de Desempenho dos Estudantes (ENADE)	37
1.6. A Provinha Brasil	39
1.7. A Prova Nacional do Concurso para o Ingresso na Carreira Docente	41
2. O que é a Teoria de Resposta ao Item?	43
2.1. Contextualização	43
2.2. O modelo logístico de três parâmetros	44
2.3. A construção da escala de proficiência	49
3. Estimação dos parâmetros dos itens e das proficiências	52
3.1. Método de estimação dos parâmetros dos itens	52
3.2. Aplicação do método de Newton-Raphson	55
3.3. Estimação das habilidades	57
4. Considerações finais	60
Referências bibliográficas	62

Prefácio

A experiência brasileira em avaliação em larga escala adquirida ao longo dos anos com a aplicação do Sistema de Avaliação da Educação Básica (SAEB) possibilitou o aprofundamento e a ampliação das técnicas utilizadas no processo para outras avaliações sistêmicas como a Prova Brasil, o Exame Nacional do Ensino Médio, o Exame Nacional para Certificação de Competências de Jovens e Adultos e o Exame Nacional de Desempenho de Estudantes.

No I Colóquio de Matemática da Região Centro-Oeste realizado em novembro de 2009, na Universidade Federal do Mato Grosso do Sul, ministrei o minicurso intitulado Os Sistemas Nacionais de Avaliação e a Teoria de Resposta ao Item (TRI). Naquela oportunidade, as mudanças implementadas pelo Ministério da Educação no Exame Nacional do Ensino Médio provocaram vários debates no meio acadêmico acerca dos benefícios de utilização da TRI nesse exame, que incluem, principalmente, a possibilidade de se comparar desempenhos de grupos de alunos que fizeram provas distintas, em momentos diferentes. Na ocasião, abordei aspectos gerais relativos aos exames nacionais de avaliação educacional e os conceitos e fundamentos básicos da TRI que dão suporte às análises dos resultados dos participantes que se submetem a esses exames.

Recentemente, foi instituída a Prova do Exame Nacional de Ingresso na Carreira Docente, que tem o objetivo principal de subsidiar os Estados, o Distrito Federal e os Municípios na realização de concursos públicos para a contratação de docentes para a educação básica. É recorrente nesses exames a elaboração prévia de uma matriz de referência, que é o cerne orientador da concepção dos itens que vão compor as provas. Em geral, as matrizes articulam, em duas ou três dimensões, a investigação concomitante das possibilidades de interação das competências esperadas para o desempenho das funções (ou dos perfis profissionais desejados) com as habilidades imprescindíveis para o desenvolvimento dessas competências e com os objetos de conhecimento por meio dos quais as habilidades são desenvolvidas. As matrizes constituem também referência para as análises de desempenho, pois permitem mensurar o grau de desenvolvimento das competências ou das habilidades pelos examinandos, além de dar uma visão ampla do profissional que se deseja avaliar/selecionar. A implantação desse novo exame nacional levanta a necessidade de se fazer um estudo comparativo dos processos de avaliação em larga escala.

Pretende-se, neste texto, estudar esses processos do ponto de vista dos múltiplos fatores envolvidos, que incluem as dimensões das matrizes de referência, os mecanismos adotados para a construção dos bancos de itens, as vantagens da aplicação da TRI, a qualidade dos instrumentos utilizados, as estratégias e situações pelas quais se obtêm informações acerca do aprendizado do avaliando e os indicadores que serão utilizados como referências para análise dos resultados e a interpretação do desempenho dos participantes.

Mauro Luiz Rabelo
Universidade de Brasília
Departamento de Matemática

Introdução

O Brasil já acumula grande experiência em avaliações educacionais de larga escala, motivada pela necessidade de se produzir indicadores que orientem políticas governamentais de melhoria da qualidade de ensino oferecido por suas instituições educacionais, distribuídas pelos mais de 5,5 mil municípios, que englobam a quinta maior população mundial, ultrapassando os 190 milhões de habitantes. Com uma área territorial superior a 8,5 milhões de km², a diversidade de realidades regionais, formas e contextos de atuação das comunidades escolares leva a um grande desafio quando se deseja introduzir metodologias de avaliação que considerem referenciais comuns para análise e estabelecimento de critérios de comparação. Alguns instrumentos balizadores da educação no Brasil foram introduzidos ao longo dos anos, referidos como Parâmetros Curriculares Nacionais (PCN) ou Diretrizes Curriculares Nacionais (DCN).

Os PCN, instrumentos de qualidade elaborados pelo Governo Federal a partir da década de 1990, estabelecem referenciais fundamentais para guiar a educação básica formal e a própria relação escola-sociedade no cotidiano. Esses instrumentos são divididos em disciplinas e abrangem práticas de organização do conteúdo, formas de abordagem e orientações de conduta a ser adotada pelos educadores em situações diversas.

Alguns marcos históricos precisam ser mencionados para contextualizar o período em que o País começou a discutir os PCN e a introduzir as avaliações de larga escala. Um deles é a própria Constituição Federal de 1988, que salvaguarda a educação como um direito fundamental, em seu Capítulo II (Dos direitos sociais) do Título II (Dos Direitos e Garantias Fundamentais), além de estabelecer em seu artigo 206 que

Art. 206 O ensino será ministrado com base nos seguintes princípios
(...)

VII - garantia de padrão de qualidade.

Oportunamente, a Lei de Diretrizes e Bases da Educação Nacional (Lei n.º 9394/1996) complementou, no Título IV (Da Organização da Educação Nacional), que

Art. 9º. A União incumbir-se-á de:
(...)

VI - assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino;

Desse modo, a União passou a ter a competência legal de assegurar processo de avaliação educacional nacional que monitorasse o requerido padrão de qualidade estabelecido na Carta Magna.

Nessa mesma época, de 5 a 9 de março de 1990, ocorria em Jomtien, na Tailândia, a Conferência Mundial sobre Educação para Todos, da qual se reproduz o inciso 1 do artigo 1 da Declaração proclamada na Conferência:

Cada pessoa — criança, jovem ou adulto — deve estar em condições de aproveitar as oportunidades educativas voltadas para satisfazer suas necessidades básicas de aprendizagem. Essas necessidades compreendem tanto os instrumentos essenciais para a aprendizagem (como a leitura e a escrita, a expressão oral, o cálculo, a solução de problemas), quanto os conteúdos básicos da aprendizagem (como conhecimentos, habilidades, valores e atitudes), necessários para que os seres humanos possam sobreviver, desenvolver plenamente suas potencialidades, viver e trabalhar com dignidade, participar plenamente do desenvolvimento, melhorar a qualidade de vida, tomar decisões fundamentadas e continuar aprendendo.

Para contemplar as diretrizes e metas do Plano de Ação da Conferência, o Brasil estabeleceu o Plano Decenal de Educação para Todos para o período 1993-2003, que consistia basicamente de um “um conjunto de diretrizes políticas voltado para a recuperação da escola fundamental no país”. É nesse contexto que surge a primeira iniciativa brasileira, em escala nacional, que objetivava conhecer o sistema educacional brasileiro em profundidade, denominada de Sistema Nacional de Avaliação da Educação Básica (SAEB). Aplicado pela primeira vez em 1990, o SAEB passou, em 1995, por uma reestruturação metodológica e incorporou as técnicas oriundas da Teoria de Resposta ao Item (TRI), exatamente pela possibilidade que essa teoria trazia de comparação dos desempenhos dos estudantes ao longo dos anos. O processo fornece dados sobre a qualidade dos sistemas educacionais do Brasil como um todo, das regiões geográficas e das unidades federadas (estados e Distrito Federal). O SAEB é feito a cada dois anos e avalia uma amostra representativa dos alunos regularmente matriculados nos 5º e 9º anos do ensino fundamental e 3º ano do ensino médio, de escolas públicas e privadas, localizadas em áreas urbana ou rural. A experiência adquirida com o uso da TRI no SAEB fez com que a metodologia a ela associada fosse difundida para outros processos de avaliação no país, abrindo perspectivas que serão objeto de análise nos capítulos a seguir.

Capítulo 1

Os Exames Nacionais de Avaliação Educacional

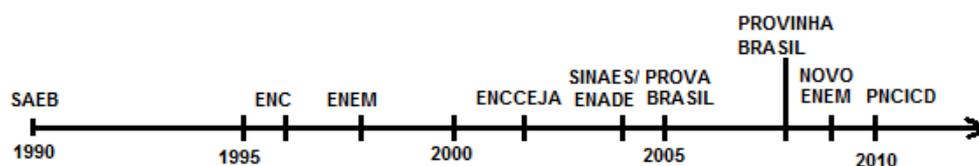
1.1 Contextualização: do SAEB à Prova Nacional de Concurso para o Ingresso na Carreira Docente

Desconsiderando as experiências com a realização dos tradicionais exames de ingresso nos cursos superiores das universidades públicas, denominados Vestibulares, será feita aqui uma comparação, do ponto de vista metodológico, dos grandes exames brasileiros que hoje estão sob a responsabilidade do Ministério da Educação, executados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Após a criação do SAEB, exclusivamente voltado para o ensino básico, o governo instituiu, em 1996, o Exame Nacional de Cursos (ENC), conhecido como Provão. Esse exame foi substituído, em 2004, pelo Exame de Desempenho de Estudantes (ENADE), que integra o Sistema Nacional de Avaliação da Educação Superior (SINAES).

Em 1998, ocorreu a primeira aplicação do Exame Nacional do Ensino Médio (ENEM), com uma perspectiva de avaliação de competências e com provas contextualizadas e interdisciplinares, que abrangiam as diversas áreas de conhecimento objeto de estudo do ensino médio. Em 2009, esse exame passou por uma reformulação metodológica e continua sendo referenciado como ENEM ou novo ENEM. Aqui será comentado apenas o exame que está hoje em vigor.

Em 2005, com o intuito de atribuir à avaliação da educação básica um caráter mais universal, criou-se a Avaliação Nacional do Rendimento Escolar, que passou a ser conhecida como Prova Brasil, que oferece *feedback* para as escolas avaliadas e cujo resultado de desempenho dos estudantes compõe um indicador de qualidade da educação denominado Índice de Desenvolvimento da Educação Básica (IDEB). Na mesma época em que a Prova Brasil foi criada, o Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA), instituído em 2002, passou por uma reformulação metodológica, que culminou com a construção de uma escala de proficiência para o exame. Em seguida, aplicou-se pela primeira vez, em 2008, a Provinha Brasil, especialmente voltada para alunos do 2.º ano da educação básica, em fase de alfabetização.

Encontra-se atualmente em discussão a implantação da Prova do Exame Nacional de Ingresso na Carreira Docente, que tem o objetivo principal de subsidiar os Estados, o Distrito Federal e os Municípios na realização de concursos públicos para a contratação de docentes para a educação básica. Este é o exame mais recente, com caráter de seleção, mas que poderá servir como processo de avaliação do perfil esperado dos profissionais que irão ingressar na carreira docente. Esse exame ainda não ganhou uma sigla, mas será aqui referenciado por suas iniciais PNCICD. Ordenando em uma escala temporal, temos a figura abaixo.



Brasil: Escala temporal dos processos nacionais de avaliação de larga escala.

1.2 O Sistema de Avaliação da Educação Básica (SAEB) e a Prova Brasil

O SAEB é o pioneiro dos sistemas nacionais de avaliação da educação hoje existentes no Brasil e contempla a aplicação de provas de língua portuguesa e matemática e questionários a uma amostra de alunos, professores e diretores. O SAEB abrange três momentos específicos do processo educacional: o fim dos anos iniciais do ensino fundamental (quinto ano), o fim do ensino fundamental (nono ano) e o fim da educação básica (terceira série do ensino médio). As análises dos resultados permitem acompanhar a evolução do desempenho e dos diversos fatores associados à qualidade e à efetividade do ensino ministrado nas escolas, inferindo-se o que os alunos sabem ou são capazes de fazer. As informações obtidas nesse processo têm o objetivo de fornecer parâmetros para a implantação de ações governamentais – tanto em nível federal quanto estadual – voltadas à correção das distorções e debilidades identificadas.

O SAEB é uma avaliação por amostra, o que significa que nem todas as turmas e estudantes das séries avaliadas participam da prova. A amostra de turmas e escolas sorteadas para participarem da avaliação é representativa das redes estadual, municipal e particular no âmbito do país, das regiões e dos estados. Dessa forma, não há resultado do SAEB por escola e por município.

Os procedimentos de amostragem de alunos utilizados baseiam-se em metodologia científica, que garante precisão nas estimativas dos parâmetros populacionais. São amostras aleatórias, probabilísticas e representativas da população de referência. De maneira geral, a população de referência do SAEB é composta pelos alunos brasileiros do ensino regular que frequentam o 5.º e 9.º ano do ensino fundamental e o 3.º ano do ensino médio, de todas as unidades da federação. Entretanto, faz-se necessário observar que a população de referência do SAEB é dinâmica, ou seja, os alunos que a constituem em cada ciclo apresentam uma estrutura variada em relação ao nível socioeconômico e cultural, fatores não controlados pelo sistema. Além disso, o próprio sistema educacional apresenta um dinamismo próprio, cujas mudanças são captadas pelo SAEB, apesar de não serem controladas por ele. Nessa linha, em cada levantamento do SAEB houve pequenas variações que deram a forma final para as populações de referência. O mesmo ocorreu na definição das amostras. Todavia, manteve-se similaridade na sua estrutura, permitindo, dessa maneira, comparações entre anos. A tabela a seguir apresenta alguns quantitativos relativos à história do SAEB.

Ciclo	Escolas	Alunos			
		4ª série EF	8ª série EF	3º ano EM	Total
1995	2.839	30.749	39.482	26.432	96.663
1997	1.933	70.445	56.490	40.261	167.196
1999	6.798	107.657	89.671	82.436	279.764
2001	6.935	114.512	100.792	72.415	287.719
2003	5.598	92.198	73.917	52.406	218.521
2005	5.940	83.929	66.353	44.540	194.822

Fonte: www.inep.gov.br

A Prova Brasil foi criada a partir do SAEB, utilizando-se a mesma matriz de referência, mas concentrando-se apenas no ensino fundamental, quinto e nono anos. A diferença fundamental entre os dois processos reside no fato de a Prova Brasil ser censitária no universo das escolas públicas urbanas, mas não alcança o ensino particular e nem o ensino médio. O quadro seguinte traz uma análise comparativa dos dois sistemas.

Quadro 1. Análise comparativa do SAEB e da Prova Brasil

Prova Brasil	SAEB
A prova foi criada e aplicada em 2005.	A primeira aplicação ocorreu em 1990.
Em 2007 e 2009, houve nova aplicação. Em 2011 está planejada sua quarta edição.	Aplica-se de dois em dois anos. A última edição foi em 2009 e a próxima será em 2011.

A Prova Brasil avalia as habilidades em Língua Portuguesa (foco em leitura) e Matemática (foco na resolução de problemas)	Alunos fazem prova de Língua Portuguesa (foco em leitura) e Matemática (foco na resolução de problemas)
Avalia apenas estudantes de ensino fundamental, de 4ª e 8ª séries.	Avalia estudantes de 4ª e 8ª séries do ensino fundamental e também estudantes da 3ª série do ensino médio.
A Prova Brasil avalia as escolas públicas localizadas em área urbana.	Avalia alunos de escolas da rede pública e de escolas privadas, localizadas nas áreas urbana e rural.
A avaliação é quase universal: todos os estudantes das séries avaliadas de todas as escolas públicas urbanas do Brasil com mais de 20 alunos na série devem fazer a prova.	A avaliação é amostral, ou seja, apenas parte dos estudantes brasileiros das séries avaliadas participam do processo.
Por ser universal, expande o alcance dos resultados oferecidos pelo SAEB. Como resultado, fornece as médias de desempenho para o Brasil, regiões e unidades da Federação, para cada um dos municípios e escolas participantes.	Por ser amostral, oferece resultados de desempenho apenas para o Brasil, regiões e unidades da Federação.
Algumas das escolas que participam da Prova Brasil ajudam a construir também os resultados do SAEB, por meio de recorte amostral.	Todos os alunos do SAEB e da Prova Brasil submetem-se a uma única avaliação.

Fonte: www.inep.gov.br (com adaptações)

Conforme se depreende da tabela acima, a partir de 2007, os dois processos passaram a ser aplicados conjuntamente, o que promoveu alterações significativas na forma de composição dos cadernos de prova do SAEB. O quadro 2 a seguir mostra a composição dos cadernos de prova do SAEB até 2005 e o quadro 3 apresenta a estrutura atual.

Quadro 2. Montagem dos cadernos – SAEB: 1995-2005

Caderno	Bloco 1	Bloco 2	Bloco 3	Caderno	Bloco 1	Bloco 2	Bloco 3
1	1	2	5	14	1	3	8
2	2	3	6	15	2	4	9
3	3	4	7	16	3	5	10
4	4	5	8	17	4	6	11
5	5	6	9	18	5	7	12
6	6	7	10	19	6	8	13
7	7	8	11	20	7	9	1
8	8	9	12	21	8	10	2
9	9	10	13	22	9	11	3
10	10	11	1	23	10	12	4
11	11	12	2	24	11	13	5
12	12	13	3	25	12	1	6
13	13	1	4	26	13	2	7

Quadro 3. Montagem dos cadernos do SAEB e Prova Brasil a partir de 2007

Cadernos	Blocos de L. Portuguesa		Blocos de Matemática	
	Posição 1	Posição 2	Posição 1	Posição 2
1	1	2	1	2
2	2	3	2	3
3	3	4	3	4
4	4	5	4	5
5	5	6	5	6
6	6	7	6	7
7	7	1	7	1
8	1	3	1	3
9	2	4	2	4
10	3	5	3	5
11	4	6	4	6
12	5	7	5	7
13	6	1	6	1
14	7	2	7	2
15	1	4	1	4
16	2	5	2	5
17	3	6	3	6
18	4	7	4	7
19	5	1	5	1
20	6	2	6	2
21	7	3	7	3

Na primeira composição, cada caderno contemplava apenas uma disciplina e a distribuição na sala era feita alternadamente: Língua Portuguesa e Matemática. Na segunda, cada caderno contempla as duas áreas, alternando-se a ordem de aparecimento de cada uma nos cadernos: o caderno 1 começa com Língua Portuguesa, o caderno 2 começa com Matemática, e prossegue-se alternando as ordens, de modo que Língua Portuguesa aparece onze vezes na primeira posição e a Matemática, dez vezes.

Essa forma de composição dos cadernos é denominada de blocos incompletos balanceados (BIB). Pode-se observar a regularidade de frequência de ocorrência de cada bloco nos cadernos, assim como a sua posição. Por exemplo, no quadro 2, o bloco 1 aparece seis vezes, sendo duas vezes em cada uma das três posições possíveis. Isso garante que as questões não sejam apresentadas aos alunos sempre na mesma posição no caderno, reduzindo a influência que a posição no caderno poderia dar no resultado de desempenho.

1.2.1 O que é avaliado nas provas

Do ponto de vista da elaboração de itens, as provas utilizadas no SAEB e na Prova Brasil são orientadas por matrizes de referência. As matrizes constituem suporte para a elaboração de itens de múltipla escolha que alimentam o banco nacional de itens (BNI), utilizado para a construção das provas, e são referências para as análises de desempenho, possibilitando a interpretação qualitativa das escalas de proficiência construídas após a aplicação dos testes.

As matrizes de referência do SAEB e da Prova Brasil contemplam as habilidades consideradas essenciais em cada etapa do ensino básico avaliadas. Elas são compostas por um conjunto de descritores que incorporam o objeto de conhecimento e a operação mental necessária para a habilidade avaliada. Esses descritores são selecionados para compor a matriz considerando-se também aquilo que é possível ser avaliado por meio de itens de múltipla escolha.

O foco em matemática é a resolução de problemas, que inclui a proposição de tarefas simples com o objetivo de avaliar se o aluno tem o domínio de padrões e técnicas escolares e se consegue associá-los a problemas rotineiros do cotidiano. A opção pelo foco na resolução de problemas está relacionada com o fato de essa metodologia possibilitar o estabelecimento de relações, o desenvolvimento de capacidades de argumentação, a validação de métodos e processos, além de estimular formas de raciocínio que incluem dedução, indução, inferência e julgamento.

As matrizes de matemática apresentam os tópicos de conteúdos matemáticos desenvolvidos no ensino fundamental e médio e especialmente selecionados para a avaliação em questão. Para cada nível avaliado, apresenta-se uma lista de descritores, agrupados nos grandes temas que constituem a formação em matemática no ensino básico: espaço e forma; grandezas e medidas; números e operações/álgebra e funções; tratamento da informação. Existe uma preocupação com o que o aluno será ou não capaz de atingir de acordo com o desempenho esperado na etapa avaliada (Rabelo, 2010).

Em um esforço de síntese, poderíamos assim resumir o que se espera que os estudantes desenvolvam em cada um desses temas à medida que se avança no percurso escolar:

Espaço e forma: inicia-se com a identificação das formas geométricas presentes em representações de objetos conhecidos, seguida da identificação de algumas características de formas geométricas planas e espaciais e das semelhanças e diferenças entre figuras geométricas planas, seus elementos e algumas relações em diferentes representações. O processo evolui para a solução de tarefas utilizando relações de congruência e semelhança de triângulos, relações métricas do triângulo retângulo e algumas propriedades de polígonos e círculos, culminando com a identificação de semelhanças e diferenças entre figuras geométricas espaciais e a solução de tarefas utilizando relações métricas e trigonométricas do triângulo retângulo em figuras planas e espaciais, incluindo noções de geometria analítica.

Grandezas e medidas: inicialmente, espera-se que o indivíduo solucione tarefas rotineiras do contexto escolar ou trazidas do cotidiano, utilizando unidades de medidas, padronizadas ou não, de comprimento, massa, capacidade, tempo e de valor. Em seguida, ele deve aprender a utilizar algumas relações entre unidades de medidas e noções de perímetro e área, para então calcular o perímetro e a área de figuras planas, estabelecendo relações entre diferentes unidades de medida. Ao final da educação básica, espera-se que ele solucione problemas que envolvam cálculo do volume e da área da superfície de alguns sólidos geométricos.

Números e operações / álgebra e funções: inicia-se com a proposta de tarefas rotineiras do contexto escolar que envolvem as operações com números naturais, seguidas da manipulação de números racionais expressos em forma de fração, decimal e percentual, para então ampliar para operações com números inteiros e racionais, incluindo procedimentos algébricos, relações de interdependência entre grandezas, linguagem algébrica na solução de equações e cálculo aproximado com radicais. O processo amplia-se, então, para situações que podem ser externas ao contexto escolar e exigem análise, avaliação e julgamento, utilizando operações com números reais, observando regularidades e relações de interdependência entre grandezas, utilizando o princípio da contagem, noções de probabilidade, fórmulas e procedimentos algébricos gerais.

Tratamento da informação: inicialmente, o indivíduo deve localizar informações explícitas apresentadas em tabelas simples e gráficos de colunas ou barras; em seguida, solucionar tarefas utilizando informações explícitas apresentadas em tabelas simples e de dupla entrada e em gráficos de colunas ou barras; prossegue, então, na comparação de informações apresentadas em tabelas e(ou) gráficos e somente depois dessa etapa espera-se que ele associe informações apresentadas em tabelas aos gráficos que as representam. No nível de conhecimento esperado no fim da educação básica, o indivíduo deve ser capaz de inferir informações implícitas a partir de dados apresentados em tabelas e gráficos.

As matrizes de língua portuguesa, com foco em leitura, têm como objetivo avaliar a capacidade de apreender o texto como construção de conhecimento em diferentes níveis de compreensão, análise e interpretação. Essas matrizes também estão organizadas sob a forma de uma lista de descritores, agrupados em seis tópicos relacionados às habilidades desenvolvidas pelos estudantes: procedimento de leitura; implicações do suporte, do gênero e(ou) enunciador na compreensão do texto; relação entre textos; coerência e coesão no processamento do texto; relações entre recursos expressivos e efeitos de sentido; variação linguística.

Essa estrutura das matrizes de língua portuguesa aliada com a forma de interpretação e análise dos resultados da avaliação permitem refletir se a) os currículos contemplam os fenômenos de variação e mudança linguística; b) a escola utiliza a educação em língua portuguesa como propulsora para o início do processo de formação plena do aluno; c) a escola está engajada na luta contra qualquer forma de exclusão social pela linguagem (Rabelo, 2011).

1.2.2 Matrizes de Referência de Matemática

Conforme foi dito anteriormente, os itens aplicados no SAEB e na Prova Brasil são elaborados a partir de matrizes especialmente criadas para esses processos. A título de completar a informação e para que o leitor compreenda melhor a descrição do que é esperado que os alunos desenvolvam em cada um dos quatro temas da área de

Matemática, apresenta-se, a seguir, as matrizes de matemática para cada uma das séries avaliadas.

Matriz de Referência – Matemática – 4ª série (5.º ano) do ensino fundamental

Descritores do Tema I. 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.

Descritores do Tema II. Grandezas e Medidas

- 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.

Descritores do Tema III. 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%, 50%, 100%).

Descritores do Tema IV. **Tratamento da Informação**

- D27 - Ler informações e dados apresentados em tabelas.
- D28 - Ler informações e dados apresentados em gráficos (particularmente em gráficos de colunas).

Matriz de Referência – Matemática – 8ª série (9.º ano) do ensino fundamental

Descritores do Tema I. **Espaço e Forma**

- D1 - Identificar a localização/movimentação de objeto, em mapas, croquis e outras representações gráficas.
- D2 - Identificar propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando-as com suas planificações.
- D3 - Identificar propriedades de triângulos pela comparação de medidas de lados e ângulos.
- D4 - Identificar relação entre quadriláteros, por meio de suas propriedades.
- 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.
- D6 - Reconhecer ângulos como mudança de direção ou giros, identificando ângulos retos e não retos.
- D7 - Reconhecer que as imagens de uma figura construída por uma transformação homotética são semelhantes, identificando propriedades e/ou medidas que se modificam ou não se alteram.
- D8 - Resolver problema utilizando a propriedade dos polígonos (soma de seus ângulos internos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares).
- D9 - Interpretar informações apresentadas por meio de coordenadas cartesianas.
- D10 - Utilizar relações métricas do triângulo retângulo para resolver problemas significativos.
- D11 - Reconhecer círculo/circunferência, seus elementos e algumas de suas relações.

Descritores do Tema II. **Grandezas e Medidas**

- D12 - Resolver problema envolvendo o cálculo de perímetro de figuras planas.
- D13 - Resolver problema envolvendo o cálculo de área de figuras planas.
- D14 - Resolver problema envolvendo noções de volume.
- D15 - Resolver problema envolvendo relações entre diferentes unidades de medida.

Descritores do Tema III. **Números e Operações /Álgebra e Funções**

- D16 - Identificar a localização de números inteiros na reta numérica.
- D17 - Identificar a localização de números racionais na reta numérica.
- D18 - Efetuar cálculos com números inteiros envolvendo as operações (adição, subtração, multiplicação, divisão e potenciação).

- D19 - Resolver problema com números naturais envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão e potenciação).
- D20 - Resolver problema com números inteiros envolvendo as operações (adição, subtração, multiplicação, divisão e potenciação).
- D21 - Reconhecer as diferentes representações de um número racional.
- D22 - Identificar fração como representação que pode estar associada a diferentes significados.
- D23 - Identificar frações equivalentes.
- D24 - Reconhecer as representações decimais dos números racionais como uma extensão do sistema de numeração decimal identificando a existência de “ordens” como décimos, centésimos e milésimos.
- D25 - Efetuar cálculos que envolvam operações com números racionais (adição, subtração, multiplicação, divisão e potenciação).
- D26 - Resolver problema com números racionais que envolvam as operações (adição, subtração, multiplicação, divisão e potenciação).
- D27 - Efetuar cálculos simples com valores aproximados de radicais.
- D28 - Resolver problema que envolva porcentagem.
- D29 - Resolver problema que envolva variações proporcionais, diretas ou inversas entre grandezas.
- D30 - Calcular o valor numérico de uma expressão algébrica.
- D31 - Resolver problema que envolva equação de segundo grau.
- D32 - Identificar a expressão algébrica que expressa uma regularidade observada em seqüências de números ou figuras (padrões).
- D33 - Identificar uma equação ou uma inequação de primeiro grau que expressa um problema.
- D34 - Identificar um sistema de equações do primeiro grau que expressa um problema.
- D35 - Identificar a relação entre as representações algébrica e geométrica de um sistema de equações de primeiro grau.

Descritores do Tema IV. **Tratamento da Informação**

- D36 - Resolver problema envolvendo informações apresentadas em tabelas e/ou gráficos.
- D37 - Associar informações apresentadas em listas e/ou tabelas simples aos gráficos que as representam e vice-versa.

Matriz de Referência – Matemática – 3.º ano do ensino médio

Descritores do Tema I. **Espaço e Forma**

- D1 - Identificar figuras semelhantes mediante o reconhecimento de relações de proporcionalidade.
- D2 - Reconhecer aplicações das relações métricas do triângulo retângulo em um problema que envolva figuras planas ou espaciais.
- D3 - Relacionar diferentes poliedros ou corpos redondos com suas planificações ou vistas.
- D4 - Identificar a relação entre o número de vértices, faces e/ou arestas de poliedros expressa em um problema.
- D5 - Resolver problema que envolva razões trigonométricas no triângulo retângulo (seno, cosseno, tangente).
- D6 - Identificar a localização de pontos no plano cartesiano.
- D7 - Interpretar geometricamente os coeficientes da equação de uma reta.
- D8 - Identificar a equação de uma reta apresentada a partir de dois pontos dados ou de um ponto e sua inclinação.
- D9 - Relacionar a determinação do ponto de interseção de duas ou mais retas com a resolução de um sistema de equações com duas incógnitas.
- D10 - Reconhecer entre as equações de 2º grau com duas incógnitas, as que representam circunferências.

Descritores do Tema II. **Grandezas e Medidas**

- D11 - Resolver problema envolvendo o cálculo de perímetro de figuras planas.
- D12 - Resolver problema envolvendo o cálculo de área de figuras planas.
- D13 - Resolver problema envolvendo a área total e/ou volume de um sólido (prisma, pirâmide, cilindro, cone, esfera).

Descritores do Tema III. **Números e Operações / Álgebra e Funções**

- D14 - Identificar a localização de números reais na reta numérica.
- D15 - Resolver problema que envolva variações proporcionais, diretas ou inversas entre grandezas.
- D16 - Resolver problema que envolva porcentagem.
- D17 - Resolver problema que envolva equação de segundo grau.
- D18 - Reconhecer expressão algébrica que representa uma função a partir de uma tabela.
- D19 - Resolver problema envolvendo uma função de primeiro grau.
- D20 - Analisar crescimento/decrescimento, zeros de funções reais apresentadas em gráficos.
- D21 - Identificar o gráfico que representa uma situação descrita em um texto.
- D22 - Resolver problema envolvendo PA/PG dada a fórmula do termo geral.
- D23 - Reconhecer o gráfico de uma função polinomial de primeiro grau por meio de seus coeficientes.
- D24 - Reconhecer a representação algébrica de uma função do primeiro grau, dado o seu gráfico.
- D25 - Resolver problemas que envolvam os pontos de máximo ou de mínimo no gráfico de uma função polinomial do segundo grau.
- D26 - Relacionar as raízes de um polinômio com sua decomposição em fatores do primeiro grau.
- D27 - Identificar a representação algébrica e/ou gráfica de uma função exponencial.
- D28 - Identificar a representação algébrica e/ou gráfica de uma função logarítmica reconhecendo-a como inversa da função exponencial.
- D29 - Resolver problema que envolva função exponencial.
- D30 - Identificar gráficos de funções trigonométricas (seno, co-seno, tangente) reconhecendo suas propriedades.
- D31 - Determinar a solução de um sistema linear associando-o a uma matriz.
- D32 - Resolver o problema de contagem utilizando o princípio multiplicativo ou noções de permutação simples e/ou combinação simples.
- D33 - Calcular a probabilidade de um evento.

Descritores do Tema IV. **Tratamento da Informação**

- D34 - Resolver problema envolvendo informações apresentadas em tabelas e/ou gráficos.
- D35 - Associar informações apresentadas em listas e/ou tabelas simples aos gráficos que as representam e vice-versa.

1.2.3 Análise do desempenho dos estudantes – aplicação da TRI

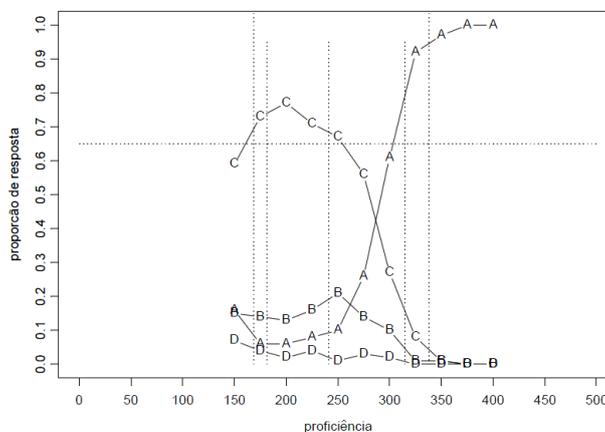
Do ponto de vista do desempenho, o SAEB tem proporcionado reflexões importantes acerca do aprendizado dos estudantes. Vamos ilustrar isso analisando o

desempenho dos estudantes em dois itens de matemática já aplicados, observando o que se esperava avaliar e o que de fato foi revelado a partir dos resultados.

Exemplo 1. O item a seguir foi aplicado aos estudantes do 9.º ano (8.ª série) para avaliar se eles seriam capazes de reconhecer as diferentes representações de um número racional. A partir das análises de desempenho, verificou-se que o item estaria inserido no nível 300 da escala de proficiência (escala com média 250 e desvio-padrão igual a 50 pontos), ou seja, somente 12,5% dos alunos 9.º ano dominam o conteúdo avaliado no item.

O número 0,25 pode ser representado pela fração

- (A) $\frac{1}{4}$
- (B) $\frac{1}{2}$
- (C) $\frac{2}{5}$
- (D) $\frac{1}{8}$



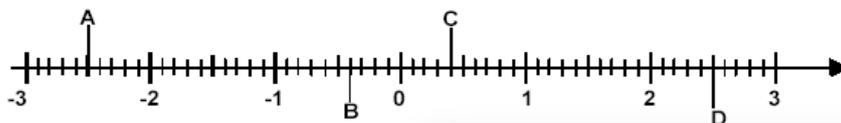
ÍNDICES							PROPORÇÕES DE RESPOSTA					
GAB	DIFI	DISCR	ABAI	ACIM	BISSE	A	B	C	D	E	""	""
A	0.27	0.59	0.07	0.66	0.72	0.27	0.14	0.56	0.03	0.00	0.01	0.00

Fonte: INEP/MEC

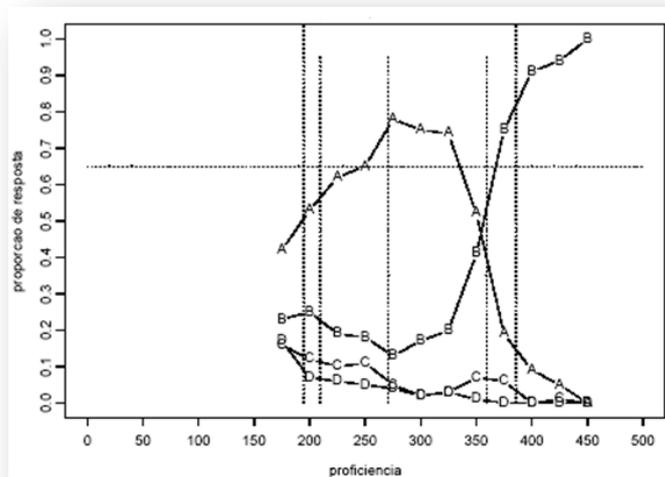
A análise dos dados acima mostra que o item revelou-se difícil para os respondentes, pois somente 27% deles acertaram. O item teve alta discriminação e alto coeficiente bisserial. A opção errada (C) atraiu 56% dos alunos, evidenciando que esses alunos não dominam os conceitos de números decimais e frações. Provavelmente, essa opção foi escolhida por conter os números 2 e 5, que aparecem no enunciado (0,25).

Exemplo 2. O item a seguir foi aplicado aos estudantes do 3.º ano do ensino médio para avaliar se eles seriam capazes de identificar a localização de números racionais na reta numérica. A partir das análises de desempenho, verificou-se que o item estaria inserido no nível 375 da escala de proficiência, de onde se conclui que apenas 6,85% dos estudantes brasileiros do fim da escolaridade básica dominam esse conteúdo.

Na reta abaixo, qual a letra que indica a localização da fração $\frac{-2}{5}$?



- (A) A
- (B) B
- (C) C
- (D) D



ÍNDICES						PROPORÇÕES DE RESPOSTA				
GAB	DIFI	DISCR	ABAI	ACIM	BISSE	A	B	C	D	E
B	0.26	0.35	0.15	0.50	0.49	0.61	0.26	0.07	0.04	0.00

Fonte: INEP/MEC

Este item exige a conversão de fração em decimal. Foi um item muito difícil para os estudantes, com apenas 26% de acerto, sendo que o percentual de acerto do grupo superior foi de 50% e o do grupo inferior 15%. A opção (A), que corresponde ao ponto -2,5, atraiu 61% dos alunos e foi a alternativa modal até o nível 350 de proficiência. É incrível que no 3.º ano do EM uma quantidade enorme de alunos ainda identifique a fração $\frac{2}{5}$ com o número 2,5.

Outros itens que avaliam o desenvolvimento dessa mesma habilidade foram aplicados para os estudantes concluintes do ensino fundamental e do ensino médio e o resultado foi praticamente o mesmo, o que nos leva a concluir que, infelizmente, a grande maioria dos estudantes brasileiros conclui o ensino médio sem saber o significado de fração.

Ainda com relação aos resultados das aplicações das provas, cabe ressaltar que as análises estatísticas clássicas, que permitem atribuir nota aos respondentes de um teste, apresentam algumas limitações, pois ficam restritas a um determinado momento, a um determinado teste e ao grupo que respondeu à prova, dificultando o acompanhamento pedagógico ao longo dos anos. Para acompanhar a evolução dos

alunos e estabelecer comparações entre séries diferentes ou mesmo entre grupos que responderam a provas diferentes, foi necessário desenvolver uma outra metodologia de construção do processo e de análise dos dados gerados após aplicação dos instrumentos. O modelo utilizado vem da psicometria moderna – a Teoria de Resposta ao Item (TRI). Pela TRI, o grau de conhecimento dos alunos é obtido por meio das características dos itens, de modo que alunos que acertam um mesmo número de itens de uma prova podem receber notas diferentes em razão de características específicas dos itens acertados. Essas características incluem a discriminação, a dificuldade e a probabilidade de acerto ao acaso, que serão explicadas em mais detalhes posteriormente.

Por meio dessas características, a TRI permite medir o nível de proficiência de cada aluno e, a partir dessa unidade de medida, construir uma escala. Essa é uma das grandes vantagens da TRI em relação à teoria clássica dos testes (TCT). Com a criação dessa escala, alunos submetidos a diferentes provas podem ser comparados entre si. Contudo, para que essa comparação seja legítima, é necessário realizar uma equalização, ou seja, é necessário transformar as medidas (parâmetros dos itens e desempenho dos alunos) para que elas fiquem em uma mesma escala. Por meio dessas análises, pode-se aferir o nível em que uma escola se encontra na escala de proficiência das disciplinas e dos níveis de ensino avaliados, no que concerne às habilidades desenvolvidas pelos alunos. Isso nos permite fazer comparações e acompanhar progressos de um sistema educacional ao longo dos anos. É importante enfatizar que as escalas de proficiência têm caráter cumulativo, ou seja, os alunos que possuem domínio em um conjunto de habilidades de um nível têm alta probabilidade de serem proficientes nas habilidades anteriores a esse nível da escala.

Com isso, compara-se o desempenho dos alunos nas séries e disciplinas avaliadas no percurso de seu desenvolvimento escolar. Com o uso dessa técnica, os resultados do SAEB e da Prova Brasil são apresentados em escalas de proficiência por disciplina, que descrevem habilidades que os alunos são capazes de demonstrar nos diversos níveis que compõem as escalas. As habilidades são cumulativas e vão se tornando mais complexas à medida que se avança na escala.

Para permitir a comparabilidade ao longo dos anos, as provas de um ano incluem alguns itens da aplicação anterior, chamados de “itens de ligação”. A partir desses itens comuns é que se faz a equalização de resultados de um ciclo para o outro, processo chamado de equalização de grupos não equivalentes com itens comuns. Na montagem

dos cadernos a partir da estrutura explicitada nos quadros 2 e 3, alguns blocos são constituídos exatamente dos itens de ligação.

Nesses dois exames, a escala de cada disciplina é numérica e varia, em tese, de 0 a 500. Escolheu-se como referência em matemática o desempenho dos alunos da 8ª série (9.º ano) do SAEB 1997, com média 250 e desvio-padrão de 50. Como os números indicam apenas uma posição, é feita uma interpretação pedagógica dos resultados por meio da descrição, em cada nível, do grupo de habilidades que os alunos demonstraram ter desenvolvido, ao responderem as questões das provas. É possível saber, pela localização numérica do desempenho na escala, quais habilidades os alunos já construíram, quais eles estão desenvolvendo e quais ainda faltam ser alcançadas. Como a escala é única para todas as séries, não é esperado, por exemplo, que alunos da 4.ª série alcancem os níveis finais da escala, pois esses representam as habilidades desenvolvidas ao longo de todo o percurso da educação básica.

Por exemplo, a média **239,5** (média nacional em matemática para o 9.º ano da rede urbana) em 2007 indica que os estudantes que estão nesse nível conseguem, entre outras ações, localizar dados em tabelas, identificar gráfico de colunas correspondentes a números positivos e negativos, converter medidas de peso e calcular o perímetro e área de figuras. Alunos com essa média também têm desenvolvidas as capacidades descritas em níveis mais baixos da escala do SAEB, como a de calcular resultados de subtrações complexas, ler horas em relógios de ponteiros e digital, estimar medida de comprimento usando unidades não convencionais e reconhecer a decomposição em dezenas e unidades de números naturais.

Nem todos os níveis das escalas são passíveis de interpretação, pois pode ocorrer de nenhum item do teste localizar-se naquele nível. O quadro 4 a seguir ilustra os níveis que são descritos de acordo com o ano.

Quadro 4. Níveis interpretados da escala de proficiência - SAEB e Prova Brasil

Nível de escolaridade	125	150	175	200	225	250	275	300	325	350	375	400	425
5º ano EF													
9º ano EF													
3º ano EM													

1.2.4 Interpretação dos níveis da escala

A partir da análise do desempenho de todos os participantes e de um novo olhar sobre todos os itens aplicados nas provas, é feita, por especialistas das duas áreas avaliadas, uma interpretação dos níveis da escala de proficiência, descrevendo-se o que significa, em termos de habilidades e conhecimentos, um indivíduo estar posicionado em uma determinada faixa da escala. A tabela seguinte, extraída do sítio www.inep.gov.br, corresponde à interpretação dos níveis da escala de matemática da Prova Brasil de 2007, para os alunos da 4.^a e 8.^a séries do ensino fundamental.

Níveis	Descrições das habilidades de acordo com o nível
125	Neste nível, os alunos da 4. ^a e da 8. ^a série resolvem problemas de cálculo de área com base na contagem das unidades de uma malha quadriculada.
150	<p>Neste nível, os alunos da 4.^a e da 8.^a série são capazes de:</p> <ul style="list-style-type: none"> ✓ resolver problemas envolvendo adição ou subtração estabelecendo relação entre diferentes unidades monetárias (representando um mesmo valor ou numa situação de troca, incluindo a representação dos valores por numerais decimais); ✓ calcular adição com números naturais de três algarismos, com reserva; ✓ reconhecer o valor posicional dos algarismos em números naturais; ✓ localizar números naturais (informados) na reta numérica; ✓ ler informações em tabela de coluna única; ✓ identificar quadriláteros.
175	<p>Os alunos das duas séries, neste nível:</p> <ul style="list-style-type: none"> ✓ identificam a localização (lateralidade) ou movimentação de objeto tomando como referência a própria posição; ✓ identificam figuras planas pelos lados e pelo ângulo reto; ✓ reconhecem a representação decimal de medida de comprimento (cm) e identificam sua localização na reta numérica; ✓ reconhecem a escrita por extenso, a composição e decomposição em dezenas e unidades, de números naturais, considerando o seu valor posicional na base decimal; ✓ efetuam multiplicação com reserva tendo por multiplicador um número com 1 algarismo; ✓ lêem informações em tabelas de dupla entrada; ✓ resolvem problemas: <ul style="list-style-type: none"> • relacionando diferentes unidades de uma mesma medida para cálculo de intervalos (dias e semanas, horas e minutos) e de comprimento (m e cm); • envolvendo soma de números naturais ou racionais na forma decimal, constituídos pelo mesmo número de casas decimais e por até três algarismos.

200	<p>Os alunos da 4ª e da 8ª séries:</p> <ul style="list-style-type: none">✓ estimam medida de comprimento usando unidades convencionais e não convencionais;✓ ler informações em gráfico de colunas;✓ reconhecer a modificação da área de uma figura pela utilização de malha quadriculada;✓ reconhecer a composição e decomposição de números naturais em sua forma polinomial e✓ localizar objeto em representação gráfica, utilizando referencial diferente da própria posição;✓ solucionam problemas:<ul style="list-style-type: none">• de subtração com números naturais de quatro algarismos, reserva e zero no minuendo;• de divisão de números naturais (dividendo: duas ordens e divisor: uma ordem). <p>Na 8ª série, os alunos são ainda capazes de resolver problemas interpretando e relacionando informações apresentadas em gráfico de colunas simples e tabela de quatro colunas.</p>
225	<p>Os alunos da 4ª e da 8ª séries apresentam as seguintes habilidades:</p> <ul style="list-style-type: none">✓ calculam divisão (divisor: duas ordens);✓ identificam propriedades comuns e diferenças em sólidos geométricos (número de faces);✓ resolvem problemas:<ul style="list-style-type: none">• de soma, a partir de dados informados em tabelas de dupla entrada;• de multiplicação de quantidades fracionadas utilizando medida de capacidade (L);• relacionando diferentes unidades de medida de tempo (mês/trimestre/ano);• utilizando noções de porcentagem ($25\% = 1/4$ e $50\% = 1/2$).✓ localizam informações em gráficos de colunas duplas;✓ calculam perímetro de figura pela soma dos lados de unidades de medida dadas em uma malha quadriculada;✓ identificam a representação numérica de uma fração dada em uma figura. <p>Na 8ª série:</p> <ul style="list-style-type: none">✓ identificam quadriláteros pelas características de seu lados e ângulos;✓ calculam o perímetro de figuras sem o apoio de malhas quadriculadas.

250

Nas duas séries, os alunos:

- ✓ calculam expressão numérica (soma e subtração) envolvendo o uso de parênteses e colchetes;
- ✓ identificam a localização de números racionais (forma decimal) na reta numérica;
- ✓ calculam e comparam as modificações do perímetro de uma figura plana desenhada em uma malha quadriculada;
- ✓ resolvem problemas:
 - de soma envolvendo combinações (três variáveis por duas);
 - lendo informações de dados apresentadas em gráficos de colunas simples, requerendo a relação inversa entre o que se pede e a altura das colunas,
 - de multiplicação (envolvendo noções de configuração retangular: fileiras x colunas)
 - realizando cálculo de conversão de medidas: de tempo (dias/anos), de temperatura (identificando sua representação numérica na forma decimal), de comprimento (m/Km) e de capacidade (ml/L).
- ✓ identificam figuras planas relacionando-as às suas nomenclaturas convencionais (triângulo, losango, pentágono e hexágono);
- ✓ identificam a representação decimal (dada em milésimos) de uma fração.

Além dessas habilidades, os alunos da 8ª série:

- ✓ identificam a movimentação de objeto em croquis utilizando conceitos de lateralidade;
- ✓ identificam entre diversas planificações aquela que permite a construção de um cubo;
- ✓ localizam números inteiros e racionais (decimais) negativos, na reta numérica;
- ✓ reconhecem e calculam ângulos retos e não-retos associando-os às suas respectivas figuras.

5

Neste nível, os alunos das duas séries:

- ✓ identificam as posições dos lados de quadriláteros (paralelismo);
- ✓ identificam a escrita numérica decimal de um inteiro fracionado (fração própria e imprópria) na reta numérica;
- ✓ identificam poliedros e corpos redondos, relacionando-os às suas planificações;
- ✓ resolvem problemas:
 - utilizando multiplicação e divisão, em situação combinatória;
 - de soma e subtração de números racionais (decimais) na forma do sistema monetário brasileiro, em situações complexas;
 - estimando medidas de grandezas utilizando unidades convencionais (L).

Na 8ª série:

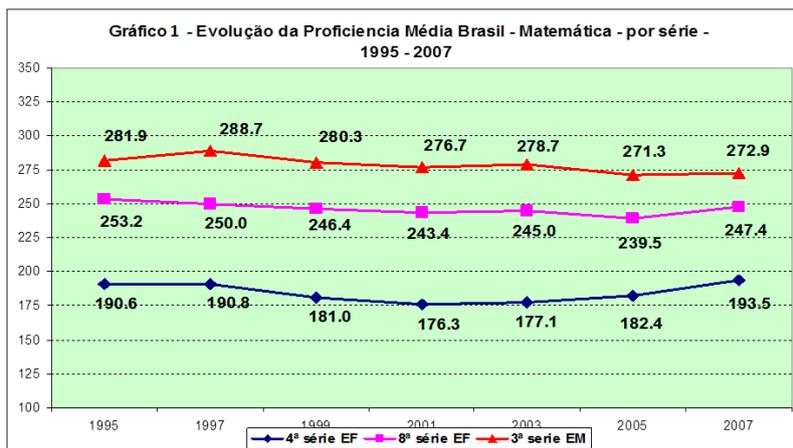
- ✓ efetuam cálculos de números inteiros positivos que requerem o reconhecimento do algoritmo da divisão inexata;
- ✓ calculam o valor numérico de uma expressão algébrica incluindo potenciação;
- ✓ identificam a localização aproximada de números inteiros não ordenados, em uma reta onde a escala não é unitária;
- ✓ resolvem situações-problema por meio da leitura e interpretação de informações dadas em percentuais em tabelas;
- ✓ solucionam problemas de cálculo de área com base em informações sobre os ângulos de uma figura e
- ✓ resolvem problemas que envolvem variação proporcional.

300	<p>Os alunos da 4ª e da 8ª séries:</p> <ul style="list-style-type: none">✓ resolvem problemas:<ul style="list-style-type: none">• identificando a lateralidade de um objeto, tendo por referência pontos com posição oposta à sua e envolvendo combinações;• realizando conversão e soma de medidas de comprimento e massa (m/km e g/kg).✓ identificam mais de uma forma de representar numericamente uma mesma fração e reconhecem frações equivalentes;✓ identificam um número natural (não informado) relacionando-o a uma demarcação na reta numérica. <p>Na 8ª série:</p> <ul style="list-style-type: none">✓ calculam o volume de sólidos a partir da medida de suas arestas;✓ identificam um número inteiro (não informado) relacionando-o a uma demarcação do início e do término de uma reta numérica;✓ identificam a equação ou inequação do 1º grau adequada para a solução de um problema;✓ solucionam problemas:<ul style="list-style-type: none">• envolvendo a propriedade dos polígonos regulares inscritos, (hexágono) para calcular o seu perímetro;• envolvendo noções de volume de figuras tridimensionais compostas por cubos, representadas num plano;• envolvendo porcentagens diversas e suas representações na forma decimal;• envolvendo o cálculo de grandezas diretamente proporcionais;• envolvendo soma de números inteiros.
325	<p>Neste nível, os alunos da 8ª série:</p> <ul style="list-style-type: none">✓ resolvem problemas:<ul style="list-style-type: none">• calculando a ampliação, redução ou conservação da medida (informada inicialmente) de ângulos, lados e área de figuras planas;• para cálculo numérico de uma expressão algébrica em sua forma fracionária;• envolvendo porcentagens diversas e suas representações na forma fracionária;• de adição e multiplicação, envolvendo identificação de um sistema de equações do primeiro grau com duas variáveis.✓ reconhecem propriedades referentes às medidas de lados e ângulos dos triângulos;✓ reconhecem as diferentes representações decimais de um número fracionário, identificando suas ordens (décimos, centésimos, milésimos);✓ efetuam cálculos de raízes quadradas;✓ realizam operações e estabelecem relações utilizando os elementos de um círculo ou circunferência (raio, diâmetro, corda).

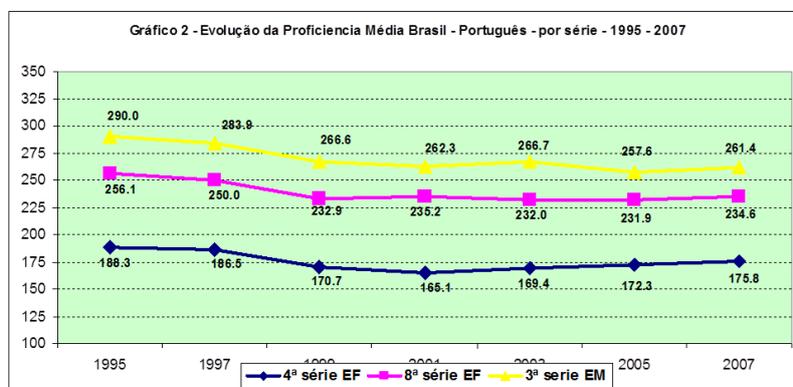
350	Além das demais habilidades demonstradas no nível anterior, neste nível, os alunos da 8ª série: <ul style="list-style-type: none">✓ reconhecem os diferentes tipos de triângulos pelas características de lados e ângulos;✓ identificam propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando as últimas às suas planificações;✓ efetuam cálculos de divisão com números racionais (forma fracionária e decimal simultaneamente);✓ resolvem problemas utilizando propriedades dos polígonos (número de diagonais, soma de ângulos internos, valor de cada ângulo interno ou externo), inclusive por meio de equação do 1º grau.
375	Neste nível, os alunos da 8ª série: <ul style="list-style-type: none">✓ interpretam informações dadas em coordenadas cartesianas;✓ efetuam uma adição de frações com denominadores diferentes;✓ resolvem problemas envolvendo círculos concêntricos;✓ solucionam problemas envolvendo a determinação das medidas da base e da altura de um triângulo traçado em malha quadriculada por contagem;✓ resolvem problemas calculando a área de triângulos por meio da aplicação do Teorema de Pitágoras;✓ calculam a área de figuras simples (triângulo, paralelogramo, retângulo, trapézio), inclusive utilizando composição e decomposição;✓ efetuam cálculos da raiz quadrada expressa na forma de expressões numéricas, envolvendo multiplicação e✓ solucionam problemas a partir da leitura e cálculo de dados oferecidos em uma tabela.

1.2.5 Construção da série histórica de desempenho

Do ponto de vista da série histórica de desempenho, em 2006, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) divulgou o relatório *Médias de desempenho do SAEB/2005 em perspectiva comparada*, que contempla os resultados de desempenho dos alunos de 4.ª e 8.ª séries do ensino fundamental (5.º e 9.º anos) e do 3.º ano do ensino médio, em língua portuguesa e matemática, na série histórica de 1995 a 2005. Os gráficos abaixo foram extraídos desse relatório, acrescentando-se as informações relativas à avaliação de 2007, já divulgadas. Os gráficos referem-se às médias de desempenho em matemática e em língua portuguesa no Brasil: urbano, rural, público (estadual e municipal) e privado.



Fonte: INEP/MEC



Fonte: INEP/MEC

A simples análise das séries longitudinais apresentadas nos gráficos acima já fornece algumas pistas acerca da evolução do desempenho dos estudantes brasileiros em língua portuguesa (leitura) e matemática.

1.2.6 O Índice de Desenvolvimento da Educação Básica (IDEB)

Atualmente, as médias de desempenho nas avaliações são utilizadas para o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB), que é o eixo do Programa de Metas Compromisso Todos pela Educação, do Plano de Desenvolvimento da Educação (PDE). Assim, a avaliação é uma das ações concretas para se aderir às metas do Compromisso e receber o apoio técnico/financeiro do MEC. O IDEB foi criado em 2007 e reúne, em um só indicador, dois conceitos igualmente importantes para a qualidade da educação: o fluxo escolar e as médias de desempenho nas

avaliações. São calculados os índices para cada unidade escolar pública, município, unidade da federação e para o país. Esse índice tem o objetivo de subsidiar a formulação e monitoramento de políticas educacionais, com vistas à melhoria da qualidade da educação. O indicador é calculado a partir dos dados relativos à aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho nas avaliações, o SAEB – para as unidades da federação e para o país, e a Prova Brasil – para os municípios.

De acordo com Fernandes, o IDEB tem como objetivo “detectar escolas ou redes de ensino cujos alunos apresentem baixo desempenho, bem como acompanhar, ao longo dos anos, a evolução do desempenho desses alunos”.

O cálculo do IDEB se dá pelo produto do desempenho e do rendimento escolar (taxa de aprovação dos alunos por série). A combinação entre fluxo e aprendizagem do IDEB expressa, em valores de 0 a 10, o andamento dos sistemas de ensino, em âmbito nacional, nas unidades da federação e municípios, calculada por meio da fórmula:

$$IDEB_{ji} = N_{ji}P_{ji}, \quad 0 \leq N_{ji} \leq 10; \quad 0 \leq P_{ji} \leq 1 \quad \text{e} \quad 0 \leq IDEB_{ji} \leq 10,$$

em que

i = ano do exame (SAEB e Prova Brasil) e do Censo Escolar;

N_{ji} = média da proficiência em Língua Portuguesa e Matemática, padronizada para um indicador entre 0 e 10, dos alunos da unidade j , obtida em determinada edição do exame realizado ao final da etapa de ensino;

P_{ji} = indicador de rendimento com base na taxa de aprovação da etapa de ensino dos alunos da unidade j .

Por exemplo, se uma escola, em uma série, obteve uma média padronizada de 7,0 na Prova Brasil, e o tempo médio de conclusão de cada série é de um ano, então seu IDEB será igual a $7,0 = 7,0 \times 1/1$. Já uma escola cuja média foi 7,0 e cujo tempo médio de conclusão de cada série é de dois anos terá um IDEB de $3,5 = 7,0 \times 1/2$.

Ainda de acordo com Fernandes,

O IDEB é o indicador objetivo para a verificação do cumprimento das metas fixadas no Termo de Adesão ao Compromisso Todos pela Educação, eixo do Plano de Desenvolvimento da Educação, do Ministério da Educação, que trata da educação básica. É nesse âmbito que se enquadra a ideia das metas intermediárias para o Ideb. A lógica é a de que para que o Brasil chegue à média 6,0 em 2021, cada sistema deve evoluir segundo pontos de partida distintos, e

com esforço maior daqueles que partem em pior situação, com um objetivo implícito de redução da desigualdade educacional.

A definição de uma meta nacional para o IDEB em 6,0 significa dizer que o país deve atingir em 2021, considerando os anos iniciais do ensino fundamental, o nível de qualidade educacional, em termos de proficiência e rendimento (taxa de aprovação), da média dos países desenvolvidos (média dos países membros da OCDE) observada atualmente. Essa comparação internacional foi possível devido a uma técnica de compatibilização entre a distribuição das proficiências observadas no PISA (*Programme for International Student Assessment*) e no SAEB.

A meta nacional norteia todo o cálculo das trajetórias intermediárias individuais do IDEB para o Brasil, unidades da Federação, municípios e escolas, a partir do compartilhamento do esforço necessário em cada esfera para que o País atinja a média almejada no período definido. Dessa forma, as metas intermediárias do IDEB, com início em 2007, foram calculadas nos âmbitos nacional, estadual, municipal e para cada escola, a cada dois anos.

O gráfico ao lado ilustra o comportamento esperado para a trajetória do IDEB ao longo dos anos. Nele estão destacados a meta estabelecida e o esforço necessário para que isso ocorra.

A compreensão sobre as questões contextuais da escola auxiliará a perceber quais os pontos que necessitam de mudanças e redirecionamento das ações institucionais para atingir a meta estabelecida. Por isso, a aplicação de questionários a alunos, professores, e diretores e a análise dos dados coletados também constituem instrumento fundamental nos sistemas de avaliação educacional.



1.3 O Exame Nacional do Ensino Médio (ENEM)

Implantado em 1998, o ENEM centra-se na avaliação individual de desempenho por competências ao final da educação básica. Tem como eixos estruturadores a interdisciplinaridade e a contextualização dos conhecimentos expressos na forma de situações-problema.

Além de ser um instrumento de aferição das competências e habilidades do indivíduo, ele também confere a cada participante parâmetro para a auto-avaliação e orientação de seu processo de formação continuada. O exame permite que o indivíduo faça análise daquilo que considera ter aprendido durante sua escolarização básica, das

suas expectativas em relação à continuidade de seus estudos e à sua inserção no mercado de trabalho, além de outras inúmeras informações extremamente valiosas para a formulação de políticas públicas no setor educacional.

A partir de 2009, agregou-se novas funcionalidades ao exame, ampliando-se o caráter de processo seletivo para acesso às instituições de educação superior, além de incorporar funções anteriormente atribuídas ao Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA) do ensino médio. Para se implantar essas mudanças, foi necessário recorrer às técnicas oriundas da TRI. A utilização dessa teoria no ENEM abriu a possibilidade de se construir uma série histórica do desempenho dos estudantes e dos egressos do ensino médio brasileiro, como é feito com os resultados do SAEB e da Prova Brasil.

Para subsidiar a reestruturação metodológica do ENEM, foi elaborada uma nova Matriz de Referência, a partir das matrizes de competências e habilidades que compõem o ENCCEJA do ensino médio. A Matriz é dividida nas quatro áreas de conhecimento que compõem o exame:

- (i) Linguagens, Códigos e suas Tecnologias (incluindo redação);
- (ii) Ciências Humanas e suas Tecnologias;
- (iii) Ciências da Natureza e suas Tecnologias; e
- (iv) Matemática e suas Tecnologias.

As cinco competências do antigo ENEM foram ampliadas para incluírem, no domínio de linguagens, o uso das línguas espanhola e inglesa, que não eram objeto de avaliação até 2009. Assume-se o pressuposto de que os conhecimentos adquiridos ao longo da escolarização deveriam possibilitar ao jovem o domínio de linguagens, a compreensão de fenômenos, o enfrentamento de situações-problema, a construção de argumentações e a elaboração de propostas. Na matriz do novo exame, essas competências correspondem aos eixos cognitivos básicos, a ações e operações mentais que todos os jovens e adultos devem desenvolver como recursos mínimos que os habilitam a enfrentar melhor o mundo que os cercam, com todas as suas responsabilidades e desafios.

Para cada uma das quatro áreas, organizou-se um conjunto de competências amplas a serem avaliadas, que, por sua vez, foram desdobradas em habilidades mais específicas, resultantes da associação dos conteúdos gerais aos cinco eixos cognitivos, totalizando 30 habilidades para cada uma das áreas. As competências de área foram

submetidas ao tratamento cognitivo das competências do sujeito do conhecimento e permitiram a definição de habilidades específicas, que estabelecem as ações ou operações que descrevem desempenhos a serem avaliados na prova. Nessa concepção, as referências de cada área descrevem as interações mais abrangentes ou complexas (nas competências) e as mais específicas (nas habilidades) entre as ações dos participantes, que são os sujeitos do conhecimento, com os objetos de conhecimento, selecionados e organizados a partir das Orientações Curriculares Nacionais para o Ensino Médio (OCNEM). Essa estrutura aproxima o novo ENEM dessas Orientações, sem abandonar o modelo de avaliação centrado no desenvolvimento de habilidades.

O foco da avaliação proposta é a análise de situação-problema, para a qual o participante deve mobilizar saberes cognitivos e conceituais (competências). A aprendizagem é destacada como referência à autonomia intelectual do sujeito ao final da educação básica, mediada pelos princípios da cidadania e do trabalho. As competências para a participação social incluem a criatividade, a capacidade de solucionar problemas, o senso crítico, o domínio de conhecimentos. Segundo Le Boterf (2003) e Zarifian (2003), competência é a capacidade do sujeito de agir diante de situações-problema inusitadas, a partir da mobilização dos recursos (cognitivos, sociais e afetivos) necessários à sua resolução. Desse modo, a competência não reside apenas nos recursos a serem acionados, mas na sua mobilização e articulação para que seja possível tomar decisões e fazer encaminhamentos adequados e úteis ao enfrentamento de situações, como a resolução de um problema ou a tomada de uma decisão. Assim, o desenvolvimento de competências pressupõe que o indivíduo não somente adquira recursos — como conhecimentos, habilidades, atitudes e valores —, mas construa, a partir deles, combinações apropriadas à ação.

De acordo com Marinho-Araújo *et all* (2010), “a abordagem por competência torna-se uma ferramenta que estabelece uma relação mais flexibilizada entre a construção do conhecimento e a transposição desse conhecimento para ações cotidianas”.

Com o novo modelo do ENEM, iniciou-se, em 2009, a construção de um banco de itens para o exame, que é submetido a um pré-teste para calibração. Não há montagem de blocos incompletos balanceados (BIB) para a aplicação, pois todos os participantes são submetidos à mesma prova, que é tornada pública imediatamente após sua aplicação. De fato, são aplicadas quatro provas, mas que diferem umas das outras

apenas pela ordenação dos itens. A linha de base da equalização dos resultados são os parâmetros originados no pré-teste. A interpretação da escala de proficiência tem sido elaborada, mas ainda não foi divulgada porque nem todas as faixas da escala possuíam itens que possibilitassem sua interpretação.

O impacto do exame na sociedade é grande, pois o número de inscritos supera os 4,5 milhões de indivíduos. A adesão ao exame como forma de substituição do tradicional vestibular tem crescido ao longo dos anos, e isso é um dos grandes atrativos, já que a escolha do curso e da instituição podem ser feitas a *posteriori* da divulgação dos resultados de desempenho, por meio de um sistema informatizado que armazena até 5 escolhas de cada candidato, para, posteriormente, alocar os pretendentes nas instituições e cursos selecionados. Recentemente, o Conselho Universitário da Universidade Federal do Rio de Janeiro, uma das maiores do país, decidiu adotar o ENEM como forma exclusiva de acesso a seus cursos de graduação a partir de 2012. Isso deverá estimular outras grandes instituições a fazer o mesmo.

1.3.1 Características do exame

A Portaria Inep/MEC n.º 109, de 27 de maio de 2009, estabeleceu a sistemática para a realização do Exame Nacional do Ensino Médio no exercício de 2009 (ENEM/2009) como procedimento de avaliação do desempenho escolar e acadêmico dos participantes, para aferir o desenvolvimento das competências e habilidades fundamentais ao exercício da cidadania. São objetivos do exame:

- I - oferecer uma referência para que cada cidadão possa proceder à sua auto-avaliação com vistas às suas escolhas futuras, tanto em relação ao mundo do trabalho quanto em relação à continuidade de estudos;
- II - estruturar uma avaliação ao final da educação básica que sirva como modalidade alternativa ou complementar aos processos de seleção nos diferentes setores do mundo do trabalho;
- III - estruturar uma avaliação ao final da educação básica que sirva como modalidade alternativa ou complementar aos exames de acesso aos cursos profissionalizantes, pós-médios e à Educação Superior;

IV - possibilitar a participação e criar condições de acesso a programas governamentais;

V - promover a certificação de jovens e adultos no nível de conclusão do ensino médio nos termos do artigo 38, §§ 1º e 2º da LDB;

VI - promover avaliação do desempenho acadêmico das escolas de ensino médio, de forma que cada unidade escolar receba o resultado global;

VII - promover avaliação do desempenho acadêmico dos estudantes ingressantes nas Instituições de Educação Superior.

Assim, o novo exame incorpora a possibilidade de certificação de jovens e adultos no nível de ensino médio e promove a avaliação do desempenho dos ingressantes nas Instituições de Educação Superior, cumprindo uma das funções que até então estava reservada para o Exame Nacional de Desempenho dos Estudantes (ENADE), no âmbito dos SINAES.

O exame compreende 4 (quatro) provas, contendo 45 (quarenta e cinco) questões objetivas de múltipla escolha, versando sobre as várias áreas de conhecimento em que se organizam as atividades pedagógicas da Educação Básica no Brasil, e uma proposta para redação. As questões são extraídas de um banco de itens calibrados, utilizando a teoria de resposta ao item. As 4 (quatro) provas são estruturadas nas áreas do conhecimento listadas anteriormente.

A proposta de reestruturação veio para evidenciar o papel que o exame já cumpria: ao longo de onze edições (1998 a 2008), a procura pelo ENEM subiu de 150 mil para mais de 4,5 milhões de inscritos, sendo que mais de 70% dos participantes afirmam que fazem a prova com o objetivo maior de chegar à faculdade. De acordo com o documento intitulado “Proposta: unificação dos processos seletivos das Instituições Federais de Ensino Superior a partir da reestruturação do Exame Nacional do Ensino Médio (ENEM)”, encaminhado pelo presidente do Inep/MEC à Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (Andifes) em 30 de março de 2009, o que se deseja é democratizar as oportunidades de concorrência às vagas federais de ensino superior por meio da unificação da seleção às vagas das IFES, utilizando uma única prova, e racionalizar a disputa por essas vagas, de forma a democratizar a participação nos processos de seleção para vagas em diferentes regiões do país.

Isso se justifica porque, de acordo com dados da Pnad/IBGE/2007, de todos os estudantes matriculados no primeiro ano do ensino superior, apenas 0,04% residem há menos de um ano no estado onde estudam. Além do mais, o motivo principal alegado pela grande maioria de estudantes para fazer o ENEM, qual seja, “entrar na faculdade / conseguir pontos para o vestibular”, comprova o caráter de processo seletivo adquirido pelo exame ao longo dos anos, de acordo com a tabela abaixo.

Motivos alegados para fazer o ENEM – Brasil – 2008

Categoria de Respostas	Frequência Absoluta	Percentual	
		Todas as categorias	Categorias válidas
Para testar meus conhecimentos / minha capacidade de raciocínio	377.306	12,92	15,37
Para entrar na faculdade / conseguir pontos para o vestibular	1.811.079	62,01	73,79
Para ter um bom emprego / saber se estou preparado(a) para o futuro profissional	255.619	8,75	10,41
Não sei	10.481	0,36	0,43
Subtotal	2.454.485	84,04	100,00
Sem informação	294.313	10,08	
Resposta inválida	10.683	0,37	
Não entregou o questionário socioeconômico	161.079	5,52	
Subtotal	466.075	15,96	
Total	2.920.560	100,00	

Fonte: Questionário socioeconômico do ENEM 2008

Como o nosso foco neste texto é a Matemática, listaremos a seguir somente a parte da matriz de referência correspondente a essa área e os eixos cognitivos, que são comuns a todas as áreas avaliadas. A matriz inclui as competências de área, as habilidades e os objetos de conhecimentos que são utilizados para a elaboração dos itens.

1.3.2 Matriz de Referência para o ENEM 2009 – Matemática e suas Tecnologias

EIXOS COGNITIVOS (comuns a todas as áreas de conhecimento)

- I **Dominar linguagens (DL):** dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica e das línguas espanhola e inglesa.
- II **Compreender fenômenos (CF):** construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos histórico-geográficos, da produção tecnológica e das manifestações artísticas.

- III **Enfrentar situações-problema (SP):** selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.
- IV **Construir argumentação (CA):** relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.
- V **Elaborar propostas (EP):** recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.

Competência de área 1 - Construir significados para os números naturais, inteiros, racionais e reais.

- H1 - Reconhecer, no contexto social, diferentes significados e representações dos números e operações - naturais, inteiros, racionais ou reais.
- H2 - Identificar padrões numéricos ou princípios de contagem.
- H3 - Resolver situação-problema envolvendo conhecimentos numéricos.
- H4 - Avaliar a razoabilidade de um resultado numérico na construção de argumentos sobre afirmações quantitativas.
- H5 - Avaliar propostas de intervenção na realidade utilizando conhecimentos numéricos.

Competência de área 2 - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.

- H6 - Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.
- H7 - Identificar características de figuras planas ou espaciais.
- H8 - Resolver situação-problema que envolva conhecimentos geométricos de espaço e forma.
- H9 - Utilizar conhecimentos geométricos de espaço e forma na seleção de argumentos propostos como solução de problemas do cotidiano.

Competência de área 3 - Construir noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.

- H10 - Identificar relações entre grandezas e unidades de medida.
- H11 - Utilizar a noção de escalas na leitura de representação de situação do cotidiano.
- H12 - Resolver situação-problema que envolva medidas de grandezas.
- H13 - Avaliar o resultado de uma medição na construção de um argumento consistente.
- H14 - Avaliar proposta de intervenção na realidade utilizando conhecimentos geométricos relacionados a grandezas e medidas.

Competência de área 4 - Construir noções de variação de grandezas para a compreensão da realidade e a solução de problemas do cotidiano.

- H15 - Identificar a relação de dependência entre grandezas.
- H16 - Resolver situação-problema envolvendo a variação de grandezas, direta ou inversamente proporcionais.
- H17 - Analisar informações envolvendo a variação de grandezas como recurso para a construção de argumentação.
- H18 - Avaliar propostas de intervenção na realidade envolvendo variação de grandezas.

Competência de área 5 - Modelar e resolver problemas que envolvem variáveis socioeconômicas ou técnico-científicas, usando representações algébricas.

- H19 - Identificar representações algébricas que expressem a relação entre grandezas.

- H20 - Interpretar gráfico cartesiano que represente relações entre grandezas.
- H21 - Resolver situação-problema cuja modelagem envolva conhecimentos algébricos.
- H22 - Utilizar conhecimentos algébricos/geométricos como recurso para a construção de argumentação.
- H23 - Avaliar propostas de intervenção na realidade utilizando conhecimentos algébricos.

Competência de área 6 - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.

- H24 - Utilizar informações expressas em gráficos ou tabelas para fazer inferências.
- H25 - Resolver problema com dados apresentados em tabelas ou gráficos.
- H26 - Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.

Competência de área 7 - Compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.

- H27 - Calcular medidas de tendência central ou de dispersão de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou em gráficos.
- H28 - Resolver situação-problema que envolva conhecimentos de estatística e probabilidade.
- H29 - Utilizar conhecimentos de estatística e probabilidade como recurso para a construção de argumentação.
- H30 - Avaliar propostas de intervenção na realidade utilizando conhecimentos de estatística e probabilidade.

1.3.3 Objetos de conhecimento para Matemática e suas Tecnologias

- **Conhecimentos numéricos:** operações em conjuntos numéricos (naturais, inteiros, racionais e reais), desigualdades, divisibilidade, fatoração, razões e proporções, porcentagem e juros, relações de dependência entre grandezas, sequências e progressões, princípios de contagem.
- **Conhecimentos geométricos:** características das figuras geométricas planas e espaciais; grandezas, unidades de medida e escalas; comprimentos, áreas e volumes; ângulos; posições de retas; simetrias de figuras planas ou espaciais; congruência e semelhança de triângulos; teorema de Tales; relações métricas nos triângulos; circunferências; trigonometria do ângulo agudo.
- **Conhecimentos de estatística e probabilidade:** representação e análise de dados; medidas de tendência central (médias, moda e mediana); desvios e variância; noções de probabilidade.
- **Conhecimentos algébricos:** gráficos e funções; funções algébricas do 1.º e do 2.º graus, polinomiais, racionais, exponenciais e logarítmicas; equações e inequações; relações no ciclo trigonométrico e funções trigonométricas.
- **Conhecimentos algébricos/geométricos:** plano cartesiano; retas; circunferências; paralelismo e perpendicularidade, sistemas de equações.

Finalmente, cabe salientar que, para cumprir as múltiplas funções do novo ENEM, o Inep/MEC precisou recorrer às técnicas oriundas da TRI. Para tentarmos entender como isso funciona na prática, vamos, no próximo capítulo, resgatar um pouco dos fundamentos dessa teoria.

1.4 O Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA)

Instituído pela Portaria MEC/INEP n.º 2.270, de 14 de agosto de 2002, como instrumento de avaliação para aferição de competências e habilidades de jovens e adultos em nível de conclusão do Ensino Fundamental e do Ensino Médio, o ENCCEJA tem por objetivos:

- I – 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;
- II – 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 9.394/1996 – Lei das Diretrizes e Bases da Educação Nacional (LDB);
- III – oferecer uma avaliação para fins de classificação na correção do fluxo escolar, nos termos do art. 24, inciso I, alínea “c”, da Lei 9.394/1996;
- IV – consolidar e divulgar um banco de dados com informações técnico-pedagógicas, metodológicas, operacionais, socioeconômicas e culturais que possa ser utilizado para a melhoria da qualidade na oferta da educação de jovens e adultos e dos procedimentos relativos ao ENCCEJA.
- V – construir um indicador qualitativo que possa ser incorporado à avaliação de políticas públicas da Educação de Jovens e Adultos.

De acordo com o art. 3º da referida Portaria, o ENCCEJA “avaliará competências e habilidades desenvolvidas por jovens e adultos 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, tendo por base a Matriz de Competências e Habilidades especialmente construída para o exame”.

Essas matrizes consideram, simultaneamente, as competências relativas às áreas do conhecimento e cinco eixos cognitivos, que expressam as competências cognitivas de jovens e adultos para a compreensão e realização de tarefas relacionadas às áreas

avaliadas. As cinco competências dos sujeitos referem-se a: domínio de linguagens, compreensão de fenômenos, enfrentamento e resolução de situações-problema, capacidade de argumentação e elaboração de propostas. A associação de cada uma das competências estabelecidas em cada área do conhecimento com os cinco eixos cognitivos resulta em trinta habilidades que são avaliadas em cada prova por meio de questões objetivas e pela produção de um texto (redação).

A adesão ao ENCCEJA é de caráter opcional e está disponível às Secretarias da Educação (estaduais ou municipais) que podem efetivá-la, formalmente, mediante assinatura de um Termo de Compromisso.

Em 2006, foi construído um banco de itens para o Exame e a TRI foi utilizada para construir uma escala de proficiência para o Exame, constituindo-se marco inicial de comparação para os anos subsequentes. A partir das mudanças implantadas no ENEM em 2009, os estudantes que desejam obter a certificação do ensino médio por conta de defasagem idade-série podem obtê-la por intermédio do ENEM. Desse modo, o ENCCEJA no Brasil agora se restringe ao nível fundamental.

O exame é aplicado também para os brasileiros residentes no exterior em nível de conclusão tanto do ensino fundamental (15 anos completos) quanto do ensino médio (18 anos completos). A oferta do ENCCEJA no exterior tem por objetivo proporcionar aos brasileiros que residem fora do país, temporariamente, a certificação equivalente ao ensino fundamental ou médio. As avaliações são elaboradas e aplicadas em língua portuguesa. Em 2011, o exame foi aplicado no Japão, nas cidades de Nagoia, Hamamatsu e Oizumi.

1.5 O Sistema Nacional de Avaliação da Educação Superior (SINAES)

O SINAES, instituído pela Lei n.º 10.861/2004, de 14 de abril de 2004, e regulamentado pela Portaria MEC/INEP 2.051/2004, surgiu da necessidade de revisão do modelo de avaliação da educação superior vigente no país até 2003. Em 2004, uma Comissão Especial de Avaliação, instituída pelo Governo Federal, apresentou proposta de mudanças para o modelo vigente no que se refere à sua metodologia, aos seus instrumentos e procedimentos e à legislação. A proposta foi amplamente discutida pelo Ministério da Educação (MEC) com as diversas entidades e representações da educação superior no país.

De acordo com o artigo 1º da Lei n.º 10.861/2004:

§ 1º O SINAES tem por finalidades a melhoria da qualidade da educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional e efetividade acadêmica e social e, especialmente, a promoção do aprofundamento dos compromissos e responsabilidades sociais das instituições de educação superior, por meio da valorização de sua missão pública, da promoção dos valores democráticos, do respeito à diferença e à diversidade, da afirmação da autonomia e da identidade institucional.

Desse modo, o objetivo do SINAES é assegurar um processo nacional de avaliação das instituições de educação superior, dos cursos de graduação e do desempenho acadêmico de seus estudantes. Nesse sentido, institucionalizou-se enquanto sistema de avaliação para a educação superior que integra, sob o princípio da globalidade, diversos instrumentos, tanto de informação quanto de avaliação. O SINAES também subsidia políticas públicas de regulação da educação superior no país, por meio de um processo avaliativo que permite acompanhamento e orientação dessa formação, especialmente no que concerne a credenciamento e renovação de credenciamento de instituições de educação superior (IES). Nessa esteira, segundo Marinho-Araújo (2004),

Como pano de fundo, o SINAES apresenta vários princípios, critérios, pressupostos e premissas que garantem sua fundamentação conceitual e política, além de justificar a operacionalização adotada no processo. Tais princípios são: *educação como um direito social e dever do Estado; valores sociais historicamente determinados; regulação e supervisão; prática social com objetivos educativos; respeito à identidade e à diversidade institucionais em um sistema diversificado; globalidade; legitimidade e continuidade.*

A concepção de avaliação do SINAES abrange três dimensões, que se constituem nos pilares do sistema: avaliação institucional, avaliação de cursos e avaliação do desempenho dos estudantes. A avaliação institucional no SINAES realiza-se a partir de uma auto-avaliação institucional e de uma avaliação externa promovida por uma equipe de especialistas capacitados para isso, com periodicidade trienal. Como meta dessa dimensão, pretende-se compreender e aprofundar os compromissos da IES, planejar ações destinadas à superação das dificuldades e ao aprimoramento das potencialidades, compartilhar compromissos e ações que visem ao desenvolvimento institucional. As outras duas dimensões, a avaliação de cursos e a de desempenho dos estudantes, incluem a aplicação de alguns instrumentos, contemplando a Avaliação das Condições de Ensino (ACE), o Exame Nacional de Desempenho dos Estudantes (ENADE) e o Censo da Educação Superior.

1.5.1 Exame Nacional de Desempenho dos Estudantes (ENADE)

No contexto do SINAES, o ENADE constitui um dos instrumentos que apresenta maiores possibilidades na concretização das articulações entre os processos avaliativos e os processos de desenvolvimento das potencialidades dos sujeitos envolvidos, em especial, os estudantes. De acordo com a legislação do SINAES, o ENADE

aferirá o desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares do respectivo curso de graduação, suas habilidades para ajustamento às exigências decorrentes da evolução do conhecimento e suas competências para compreender temas exteriores ao âmbito específico de sua profissão, ligados à realidade brasileira e mundial e a outras áreas do conhecimento. (Lei nº 10.861/2004, art. 5º).

Ainda, de acordo com a Portaria 2051/2004, a avaliação do desenvolvimento de conhecimentos e competências no ENADE está concebida para investigar, no curso do processo pedagógico, as habilidades, os conhecimentos e os saberes fundamentais que compõem o perfil profissional; tal processo deverá considerar as Diretrizes Curriculares, a articulação teoria e prática e o modo como as competências foram se construindo, articuladas às subjetividades individuais e sociais próprias às relações e aos contextos vivenciados. Desse modo, considera-se que o foco do ENADE é o perfil profissional que se deseja formar e a trajetória acadêmica da formação dos estudantes. A ênfase da avaliação inclui o desenvolvimento de competências necessárias ao aprofundamento da formação profissional e à compreensão, integrada, crítica e contextualizada, de temas diversos àqueles específicos da formação.

Não existe matriz de referência estabelecida para as provas do ENADE. As provas são construídas a partir das Diretrizes Curriculares estabelecidas para cada curso e contemplam 10 questões de formação geral – 8 objetivas e 2 discursivas –, comuns a todos os cursos avaliados, e 30 de conhecimentos específicos – 8 objetivas e 2 discursivas. A aplicação para cada curso é em ciclos de três anos. Iniciou-se o processo em 2004 com grupos amostrais, representativos de cada instituição/curso, que se encontram em momentos distintos de sua graduação (grupos que correspondem ao final do primeiro, denominados ingressantes, e do último ano do curso, denominados concluintes), permitindo, assim, um acompanhamento longitudinal do processo de ensino e aprendizagem. Recentemente, ampliou-se a participação para o universo de

estudantes concluintes. Para os estudantes selecionados, a participação é obrigatória, pois, o não comparecimento impede a emissão do diploma do curso superior.

A metodologia contempla a elaboração de prova única para cada curso, com itens não calibrados e que não são pré-testados. Utiliza-se apenas a Teoria Clássica dos Testes para as análises e não se constrói interpretação de escala de proficiência. A metodologia adotada também não permite a construção de série histórica para análise de evolução dos processos de aprendizagem, já que os resultados não são comparáveis. Os participantes respondem a um questionário socioeconômico, que tem por objetivos traçar o perfil dos estudantes, ingressantes e concluintes, dos cursos de graduação do país; conhecer a opinião dos estudantes a respeito do ambiente acadêmico em que realizam a sua formação; e consolidar informações para promover a melhoria das condições de ensino e dos procedimentos didático-pedagógicos.

Vários indicadores foram associados aos instrumentos que são utilizados no SINAES. O Indicador de Diferença Entre os Desempenhos Observado e Esperado (IDD) tem o propósito de trazer às instituições informações comparativas dos desempenhos de seus estudantes concluintes em relação aos resultados obtidos, em média, pelas demais instituições cujos perfis de seus estudantes ingressantes são semelhantes. Considera-se que essas informações são boas aproximações do que seria considerado o efeito do curso. O IDD é a diferença entre o desempenho médio do concluinte de um curso e o desempenho médio estimado para os concluintes desse mesmo curso e representa, portanto, quanto cada curso se destaca da média, podendo ficar acima ou abaixo do que seria esperado para ele baseando-se no perfil de seus estudantes. Em 2007, o INEP criou um indicador de qualidade das instituições de educação superior, o Índice Geral de Cursos da Instituição (IGC), que considera, em sua composição, a qualidade dos cursos de graduação e de pós-graduação (mestrado e doutorado). No que se refere à graduação, é utilizado o CPC (conceito preliminar de curso) e, no que se refere à pós-graduação, é utilizada a Nota CAPES (INEP/MEC, 2008).

O CPC é uma média de diferentes medidas da qualidade de um curso. As medidas utilizadas são: o Conceito ENADE (que mede o desempenho dos concluintes), o desempenho dos ingressantes no ENADE, o Conceito IDD e as variáveis de insumo. O dado referente às variáveis de insumo – que considera corpo docente, infraestrutura e

programa pedagógico – é formado com informações do Censo da Educação Superior e das respostas dos estudantes ao questionário socioeconômico.

A forma do cálculo do CPC tem implicações sobre a representatividade do IGC. Para um curso ter CPC é necessário que ele tenha participado do ENADE com alunos ingressantes e alunos concluintes.

A Avaliação dos Programas de Pós-graduação, realizada pela CAPES, compreende a realização do acompanhamento anual e da avaliação trienal do desempenho de todos os programas e cursos que integram o Sistema Nacional de Pós-graduação, SNPG. Os resultados desse processo, expressos pela atribuição de uma nota na escala de 1 a 7, fundamentam a deliberação CNE/MEC sobre quais cursos obterão a renovação de reconhecimento, a vigorar no triênio subsequente. A medida de qualidade da pós-graduação que compõe o IGC é uma conversão das notas fixadas pela CAPES. (INEP/MEC, 2008)

1.6 A Provinha Brasil

A Provinha Brasil é uma avaliação diagnóstica do nível de alfabetização das crianças matriculadas no segundo ano de escolarização das escolas públicas brasileiras. Essa avaliação ocorre em duas etapas, uma no início e outra ao término do ano letivo. A aplicação em períodos distintos possibilita aos professores e gestores educacionais a realização de um diagnóstico mais preciso que permite conhecer o que foi agregado na aprendizagem das crianças, em termos de habilidades de leitura e matemática no período avaliado.

Em abril de 2008, foi aplicada a 1ª edição dessa avaliação. Cerca de 3.133 municípios e 22 unidades federativas receberam o material impresso e as demais secretarias de educação puderam fazer o *download* do material na página do INEP. A partir do 2º semestre de 2008, além da disponibilização do material na *Internet*, todas as secretarias de educação do país passaram a receber o material impresso. (INEP/MEC, 2011)

A Provinha Brasil é instrumento pedagógico, sem finalidades classificatórias, que tem como objetivos principais avaliar o nível de alfabetização dos alunos/turma nos anos iniciais do ensino fundamental e diagnosticar possíveis insuficiências das habilidades de leitura, matemática e escrita. Segundo o INEP, esses objetivos

possibilitam, entre outras ações, o estabelecimento de metas pedagógicas para a rede de ensino; o planejamento de cursos de formação continuada para os professores; o investimento em medidas que garantam melhor aprendizado; o desenvolvimento de ações imediatas para a correção de possíveis distorções verificadas; a melhoria da qualidade e redução da desigualdade de ensino. O delineamento e a construção dessa avaliação prevê, sobretudo, a utilização dos resultados obtidos nas intervenções pedagógicas e gerenciais com vistas à melhoria da qualidade da alfabetização.

Até 2010, cada teste da Provinha Brasil era composto por 24 questões de múltipla escolha, com quatro opções de resposta cada uma. A partir de 2011, os testes serão compostos de 20 questões. Algumas dessas questões são lidas pelo aplicador da prova – na íntegra ou em parte – e outras são lidas apenas pelos alunos.

Em cada ano, ocorre um novo ciclo de avaliação da Provinha Brasil. Cada ciclo é composto por duas etapas. A Provinha Brasil é realizada em dois momentos durante o ano letivo: no início do 2º ano de escolarização e no final desse mesmo ano letivo. Sugere-se que o Teste 1 seja aplicado, preferencialmente, até o mês de abril, e o Teste 2, até o final de novembro.

As habilidades avaliadas por meio da Provinha Brasil estão organizadas na Matriz de Referência para Avaliação da Alfabetização e do Letramento Inicial e na Matriz de Referência para Avaliação da Alfabetização Matemática Inicial.

Como nem todas as habilidades a serem desenvolvidas durante o processo de alfabetização e letramento, como a oralidade, são passíveis de verificação em uma prova objetiva, foram selecionadas as habilidades consideradas essenciais para a alfabetização e letramento. A Matriz de Letramento é dividida em dois eixos. O primeiro é a “apropriação do sistema de escrita: habilidades relacionadas à identificação e ao reconhecimento de princípios do sistema de escrita” e o segundo é a “leitura”.

A Matriz de Matemática traz as competências e habilidades esperadas nesse nível de escolaridade, agrupadas em quatro eixos: números e operações; geometria; grandezas e medidas; e tratamento da informação.

Cada questão do teste avalia, de forma preponderante, uma das habilidades descritas em cada matriz. Assim com nas demais avaliações operacionalizadas pelo INEP, a matriz é apenas uma referência para a construção dos testes, não se aproximando de uma proposta curricular ou programa de ensino, que são bem mais amplos e complexos.

1.7 A Prova Nacional do Concurso para o Ingresso na Carreira Docente

A Portaria n.º 14, de 21 de maio de 2010, instituiu o Exame Nacional de Ingresso na Carreira Docente, que passou a ser denominado de Prova Nacional de Concurso para o Ingresso na Carreira Docente, a partir da Portaria Normativa n.º 3, de 2 de março de 2011. Essa prova tem o objetivo principal de subsidiar os Estados, o Distrito Federal e os Municípios na realização de concursos públicos para a contratação de docentes para a educação básica. A Prova tem, ainda, os objetivos de proporcionar parâmetros para a auto-avaliação dos candidatos à docência e oferecer informações para subsidiar a formulação e a avaliação das políticas públicas de formação inicial e continuada de docentes.

No âmbito da educação, a realização periódica de concursos públicos pode ser uma iniciativa de grande valia para valorização do magistério, pois oferece oportunidades iguais a todos os candidatos, permite a seleção dos profissionais com maior potencial para o exercício da carreira e promove a efetivação de professores nos quadros estaduais e municipais, assegurando-lhes um contrato de trabalho permanente. Esses fatores, por sua vez, permitem dotar as escolas de um corpo docente devidamente preparado e com as condições necessárias para desenvolver um trabalho com continuidade – aspectos tão caros à qualidade dos processos de ensino e aprendizagem. (INEP, 2011)

Como a elaboração e a aplicação do exame estarão a cargo do INEP, as secretarias que fizerem adesão ao exame poderão contratar, com maior regularidade, professores para cargos permanentes, resolvendo o problema crônico de professores temporários, denominados “precarizados”, hoje existentes na maioria das escolas públicas. Para os candidatos, a existência de uma única prova aplicada em todo o país e utilizada por diversos estados e municípios facilitará sua escolha do local de trabalho para se efetivar com professor em uma rede de ensino pública.

A Prova será aplicada anualmente, de forma descentralizada em todo o país. Cada ente federativo poderá decidir pela adesão e pela forma de utilização dos resultados: se como única prova, seguida da análise de títulos, ou como primeira prova, seguida de uma prova adicional do próprio município, por exemplo.

A iniciativa também permitirá que as provas para a seleção de professores para a educação básica tenham melhor qualidade. Isso porque a elaboração dos itens que vão compor os testes será feita com base em uma matriz de referência construída após

amplo processo de discussão. A matriz teve como ponto de partida a pesquisa e a reflexão sobre o perfil desejado para um ingressante na carreira docente no Brasil. O processo é uma iniciativa do INEP no âmbito das políticas do Ministério da Educação no sentido da valorização dos profissionais do magistério no país.

Capítulo 2

O que é a Teoria de Resposta ao Item?

2.1 Contextualização

A TRI surgiu entre os anos 50 e 60 do século passado para responder a indagações relativas aos testes de inteligência, cujos resultados variavam em função dos instrumentos de medida utilizados. Qual seria então o resultado correto? O problema era mais profundo já que “o objeto medido, a inteligência no caso, afeta diretamente o instrumento utilizado; aliás, ela é definida pelo próprio instrumento utilizado” (Pasquali, 2007). Já pensou se o metro medisse de forma diferente se estivéssemos calculando a largura de um livro ou o comprimento de uma mesa? A solução dada pela TRI a esse problema – independência do instrumento de medida em relação ao objeto que se deseja medir – utilizava modelos e algoritmos matemáticos difíceis de serem operacionalizados à época. Por isso, somente após o avanço tecnológico dos anos 80, com o desenvolvimento de *softwares* para uso prático dos algoritmos complexos que o modelo contém, essa teoria começou a ser difundida.

A utilização da TRI para análise de testes de conhecimento veio para sanar algumas limitações da Teoria Clássica dos Testes (TCT), principalmente no que diz respeito a discriminação dos itens, fidedignidade dos testes e comparabilidade de desempenho de indivíduos que se submetem a testes diferentes. Tomando as edições do ENEM já ocorridas desde 1998, é comum ouvirmos inferências comparativas dos resultados de um ano com o seguinte, simplesmente observando-se a média dos alunos na prova. Essas comparações são completamente desprovidas de fundamento e não podem ser feitas. Os resultados de um ano para o outro não são comparáveis. Com a utilização da TRI no ENEM, poderão ser tiradas conclusões interessantes e cientificamente bem fundamentadas sobre o desempenho dos estudantes e dos egressos do ensino médio brasileiro ao longo do tempo.

Uma escala de proficiência para cada uma das quatro áreas avaliadas será construída, e essa será a “régua” que vai ser utilizada para se comparar os resultados ao longo dos anos. Será possível fazer conclusões análogas ao que hoje já é inferido a partir dos dados obtidos com o SAEB e a Prova Brasil. Conforme comentamos no capítulo anterior, até 1993, o SAEB utilizou a TCT para a construção dos instrumentos,

atribuição dos escores e análise dos resultados, não havendo planejamento para uma comparação dos resultados. A partir de 1995, foi introduzido o uso da TRI para a construção de instrumentos, a atribuição de escores e a análise, de forma a viabilizar a comparação dos resultados, com a construção das escalas de proficiência. A experiência de utilização dessa metodologia no Brasil já tem, portanto, cerca de 15 anos.

2.2 O modelo logístico de três parâmetros

A TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma resposta a um item como função dos parâmetros do item e da(s) habilidade(s) do respondente. Os modelos relacionam variáveis observáveis – respostas aos itens de um teste, por exemplo – com aptidões² não observáveis e que são responsáveis pelas respostas dadas pelo indivíduo. De acordo com essa relação, quanto maior a habilidade, maior a probabilidade de acerto no item.

O que se deseja é estimar o nível de aptidão, o traço latente, do indivíduo a partir da análise das respostas dadas por ele a um conjunto de questões ou itens. Para isso, recorre-se a modelos matemáticos que relacionam as variáveis envolvidas nessa situação. O modelo utilizado para análise no ENEM é o modelo logístico de 3 parâmetros, que permite que seja estimado o nível de aptidão (ou traço latente) do respondente a partir de uma relação que fornece a probabilidade de um indivíduo acertar um item em função de sua habilidade (θ), da dificuldade, da discriminação e da probabilidade de acerto ao acaso, popularmente conhecida como “parâmetro de chute”.

Essa teoria apresenta pressupostos que devem ser observados pelos especialistas que elaboram os itens – a unidimensionalidade e a independência local – e sobre os quais vamos fazer apenas um breve comentário. A unidimensionalidade é a hipótese segundo a qual “há apenas uma aptidão dominante responsável pelo desempenho em um conjunto de itens de um teste” (Pasquali, 2003). Já a hipótese de independência local assume que “para uma dada habilidade, as respostas aos diferentes itens da prova são independentes” (Andrade & cols, 2000), ou seja, “mantidas constantes as aptidões que afetam o teste, o desempenho de um indivíduo em um item não interfere no seu desempenho em outro item” (Pasquali, 2003). Isso não quer dizer que os itens não

² Neste texto, vamos usar indistintamente as palavras aptidão, habilidade e proficiência.

possam estar correlacionados, mas que cada indivíduo dá respostas independentes para cada item do teste.

Quando um indivíduo responde a uma sequência de itens, produz um padrão de respostas, composto de acertos (valor = 1) e erros (valor = 0). Isso pode ser esquematizado em uma tabela cujas linhas são os padrões de respostas dos indivíduos submetidos ao teste e cujas colunas são os itens do teste, numerados, por exemplo, de 1 a 45 (o novo ENEM é composto de quatro provas de 45 itens cada uma). Essa tabela de “uns” e “zeros” coleciona as respostas dadas por todos os indivíduos a todos os itens de cada prova. Se forem 4,5 milhões de participantes, será, para cada área avaliada, uma tabela de 4,5 milhões de linhas por 45 colunas, chamada de matriz de padrão de respostas dos indivíduos. Somente com os recursos atuais de informática podemos manusear tabelas tão grandes.

Na TCT, a aptidão de um indivíduo que respondeu ao teste é simplesmente expressa pelo número de itens que ele acertou. Compara-se seu padrão de respostas com o gabarito e calcula-se o escore bruto fazendo-se a soma dos acertos. No caso da TRI, deseja-se descobrir qual o valor do traço latente (da habilidade) do indivíduo que explica o acerto ou o erro em cada item individualmente.

Para descobrir isso, a pergunta inicial que a TRI faz sobre o item é: qual é a probabilidade de um dado indivíduo acertar um item específico? A resposta a essa pergunta depende do nível de aptidão do indivíduo (θ) e das características do item – dificuldade (b), discriminação (a) e acerto ao acaso (c).

Entre os modelos propostos pela TRI, o modelo logístico unidimensional de 3 parâmetros é mais utilizado e representa a probabilidade de um indivíduo j responder corretamente um item i , sendo definido por:

$$P(X_{ij} = 1 | \theta_j) = c_i + \frac{(1 - c_i)}{1 + \exp[-Da_i(\theta_j - b_i)]},$$

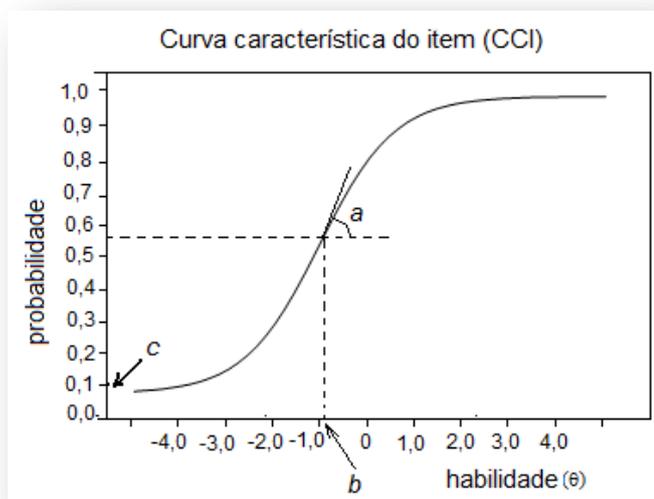
em que

- X_{ij} é a resposta ao item i (igual a 1, se o indivíduo responde corretamente, e igual a 0, caso contrário);
- $a_i > 0$ é o parâmetro de discriminação do item i ,
- b_i é o parâmetro de posição (ou de dificuldade) do item, medido na mesma escala da habilidade;

- $0 < c_i < 1$ é o parâmetro da assíntota inferior do item i , refletindo as chances de um estudante de proficiência muito baixa selecionar a opção de resposta correta;
- θ_j representa a habilidade ou traço latente do j -ésimo indivíduo;
- D é um fator de escala, que é igual a 1 na métrica logística e igual a 1,7 na métrica normal.

O número $P(X_{ij} = 1 | \theta_j)$ pode ser identificado com a proporção de respostas corretas ao item i no grupo de indivíduos com habilidade θ_j . A relação expressa pela equação acima está descrita pela curva da figura abaixo, chamada de Curva Característica do Item (CCI).

Logo, na TRI, estuda-se o comportamento do indivíduo frente a cada item que ele responde. Por isso, a base desse modelo matemático é a CCI, gráfico da função $P(\theta)$, que fornece a probabilidade de um aluno com habilidade θ responder corretamente o item do teste que está sendo analisado. A função $P(\theta)$ assume valores que vão de 0 a 1 (correspondentes a 0% a 100%) e se configura em uma curva na forma de S. Na curva, estão indicados os parâmetros a , b , c do item, que serão explicados a seguir. Os valores da habilidade (θ), no eixo horizontal, estão na escala de média igual a zero e desvio-padrão 1.



O acerto ao acaso representa as respostas dadas por “chute”. Isso ocorre principalmente com os itens que são mais difíceis, para os quais indivíduos de baixa aptidão não conhecem a resposta correta, mas arriscam qualquer uma. A TRI estima o acerto ao acaso por meio do parâmetro c , que representa a probabilidade de um aluno com baixa habilidade responder corretamente o item. Assim, se não fosse permitido “chutar”, c seria igual a zero. Em geral, são recomendados valores iguais ou inferiores a 0,20 para itens com 5 opções de resposta, como é o caso do ENEM.

A dificuldade é o valor da aptidão (θ) que é necessário para se obter uma probabilidade de acerto igual a $(1+c)/2$. Quando não é permitido o “chute”, a dificuldade é o valor da habilidade necessário para se obter uma probabilidade de 50% de acertar o item. Quanto maior for o nível de aptidão necessário para o indivíduo acertar o item, maior será a dificuldade desse item. A dificuldade é representada em uma escala padronizada, que, teoricamente, varia de $-\infty$ a $+\infty$. Na prática, seus valores situam-se tipicamente entre -3 e 3, pois, entre esses extremos, estão mais de 99,7% dos casos. Os itens cujo parâmetro b está próximo de -3 são considerados fáceis e os de parâmetro próximo de +3, difíceis. Itens cujo valor de b cai fora do intervalo (-3, 3) sugerem problemas de concepção e são, normalmente, excluídos das análises.

O nível de dificuldade ideal para os itens de um teste depende da sua finalidade. Em avaliação educacional, recomenda-se “uma distribuição de níveis de dificuldade de itens no teste dentro de uma curva normal: 10% dos itens em cada uma das duas faixas extremas, 20% em cada uma das faixas seguintes e 40% na faixa média” (Pasquali, 2003).

A discriminação é concebida como a capacidade do item de diferenciar indivíduos com diferentes habilidades. Na TRI, é definida como o poder do item para diferenciar indivíduos com magnitudes próximas da habilidade que está sendo aferida. Na CCI, a discriminação é identificada como um valor proporcional à inclinação da curva no ponto onde ela muda de concavidade (ponto de inflexão). Valores negativos da discriminação indicam que o item se comporta de uma maneira estranha, pois estariam indicando que a probabilidade de acerto do item diminui com o aumento da aptidão do sujeito. Obviamente, tais itens devem ser descartados.

A CCI evidencia o poder de discriminação do item em termos da inclinação da curva, isto é, quanto mais elevados forem os valores do parâmetro a , mais íngreme será a curva no ponto de inflexão e mais discriminativo será o item. Valores do parâmetro a próximos a zero indicam que o item tem pouco poder de discriminação, ou seja, indivíduos com habilidades muito diferentes têm, aproximadamente, a mesma probabilidade de responder corretamente ao item. Nesse caso, a CCI não seria em forma de S, mas próxima a um segmento horizontal. Um item com essa característica não cumpriria os propósitos de uma avaliação educacional, já que não discriminaria os respondentes que sabem daqueles que não sabem o que está sendo nele avaliado.

Vamos ilustrar a situação teórica descrita acima em um caso real, para um item aplicado no ENEM de 2007.

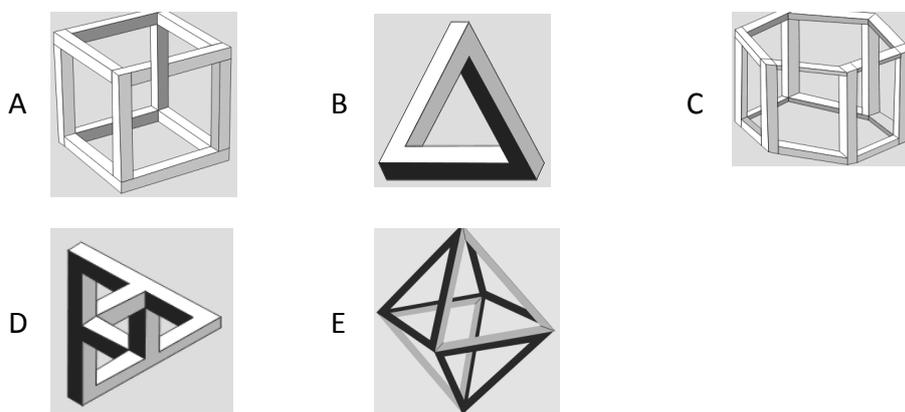
Exemplo.

Representar objetos tridimensionais em uma folha de papel nem sempre é tarefa fácil. O artista holandês Escher (1898-1972) explorou essa dificuldade criando várias figuras planas impossíveis de serem construídas como objetos tridimensionais, a exemplo da litografia **Belvedere**, reproduzida ao lado.

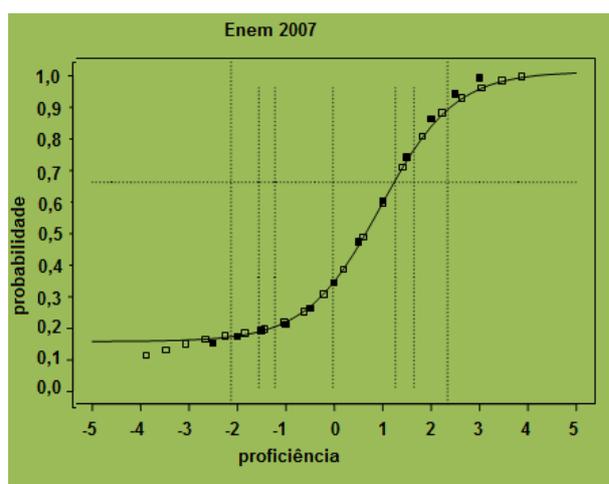
Considere que um marceneiro tenha encontrado algumas figuras supostamente desenhadas por Escher e deseje construir uma delas com ripas rígidas de madeira que tenham o mesmo tamanho.



Qual dos desenhos a seguir ele poderia reproduzir em um modelo tridimensional real?



A curva característica desse item está ilustrada abaixo. Os parâmetros do item, obtidos utilizando-se a TRI, são $a = 1,31$, $b = 0,96$ e $c = 0,15$.



Fonte: Inep/MEC

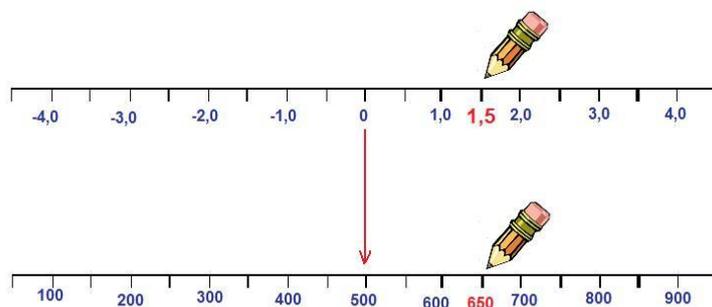
A forma da curva já diz bastante sobre o comportamento desse item, que apresentou bons parâmetros: discriminou muito bem ($a = 1,31$), está na faixa de média para alta dificuldade ($b = 0,96$) e o parâmetro de acerto ao acaso ficou abaixo do máximo permitido ($c = 0,15$).

Para responder ao item, o candidato deveria, por meio de uma habilidade de percepção espacial adequada, analisar as figuras tridimensionais dadas e verificar em qual delas as relações de visibilidade e sombreamento entre as peças de madeira correspondiam a um objeto real.

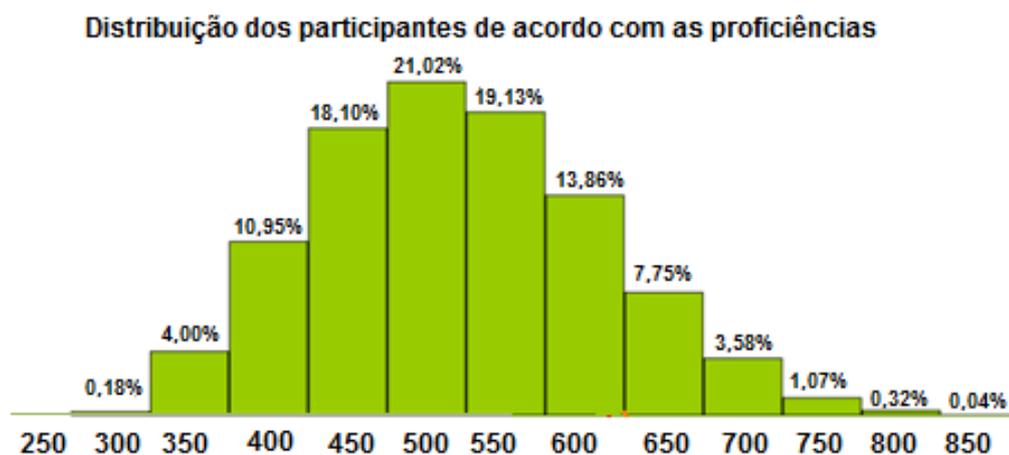
2.3 A construção da escala de proficiência

Na TRI, a habilidade pode, teoricamente, assumir qualquer valor entre $-\infty$ a $+\infty$. Assim, precisa-se estabelecer uma origem e uma unidade de medida para a definição da escala. Esses valores são escolhidos de modo a representar, respectivamente, o valor médio e o desvio-padrão das habilidades dos indivíduos que responderam o teste. Para os gráficos mostrados anteriormente, utilizou-se a escala com média igual a 0 e desvio-padrão igual a 1, comumente representada por escala (0,1). Essa escala é bastante utilizada e, nesse caso, os valores do parâmetro b variam tipicamente entre -3 e 3. Com relação ao parâmetro a , espera-se valores entre 0 e 2, sendo que os valores mais apropriados seriam aqueles maiores que 1. (Andrade & cols, 2000).

Apesar da frequente utilização da escala (0,1), o importante, em termos práticos, é a relação de ordem existente entre os pontos da escala. Por exemplo, um indivíduo com habilidade 1,5 está a 1,5 desvio-padrão acima da habilidade média. Se a escala for com média 500 e desvio-padrão 100, como atualmente se usa no ENEM, esse indivíduo teria habilidade igual a 650 nessa escala, estando a exatamente 1,5 desvio-padrão acima da média. Veja essa situação ilustrada na figura abaixo.



Os valores dos parâmetros dos itens e das habilidades dos indivíduos são obtidos utilizando-se técnicas estatísticas e matemáticas de estimação, com o auxílio de *softwares* específicos, que manipulam tabelas (matrizes) gigantescas como a matriz de padrão de respostas anteriormente mencionada. De qualquer modo, em uma população tão grande quanto a que se submete ao ENEM ou à Prova Brasil, a distribuição dos indivíduos por nível de habilidade é praticamente normal, conforme ilustra hipoteticamente a figura a seguir.



Ao receber o boletim de desempenho, o participante do ENEM se localiza nessa distribuição, de acordo com o seu desempenho no teste e seu nível de aptidão em cada uma das quatro áreas avaliadas no exame: Linguagens, Códigos e suas Tecnologias (incluindo redação); Ciências Humanas e suas Tecnologias; Ciências da Natureza e suas Tecnologias; e Matemática e suas Tecnologias. De posse das informações fornecidas, o participante pode fazer uma análise do seu nível de aptidão em cada uma dessas quatro áreas, comparativamente aos demais candidatos.

No entanto, saber que em uma escala de 0 a 1.000, o estudante obteve 550 pontos ou que, no 3.º ano do ensino médio, os alunos têm habilidade média de 400 pontos em matemática, já nos fornece alguma informação quantitativa do desempenho, mas não nos diz nada em termos qualitativos. Afinal, o que eles sabem de conteúdo? E o que eles demonstraram saber a mais em relação ao que sabiam ao concluírem o 9.º ano do ensino fundamental?

Para atribuir significado prático aos valores obtidos na escala, utiliza-se também a TRI, em um processo denominado de interpretação pedagógica das escalas de proficiência. A metodologia para essa interpretação inclui dois procedimentos básicos: identificação de “itens âncoras” de cada nível da escala e a apresentação desses itens a um painel de especialistas.

No SAEB, utilizou-se a metodologia descrita em Beaton & Allen (1992). Nessa metodologia, itens âncora são itens que caracterizam os pontos ou níveis das escalas no sentido de que a grande maioria dos avaliados situados em determinado nível acerta o item, enquanto menos da metade dos alunos situados no nível imediatamente inferior também acerta. Por exemplo, para o item ser considerado âncora no nível 400, ele deve satisfazer ao seguinte critério:

- 65% ou mais dos respondentes em torno do nível 400 acertam o item;
- menos de 50% dos alunos posicionados no nível anterior acertam o item; e
- a diferença entre os percentuais dos que acertaram seja maior que 30%.

Este critério é bastante rígido, e, por isso, uma modificação tem sido empregada:

- 65% ou mais dos respondentes em torno do nível 400 acertam o item;
- menos de 65% dos alunos posicionados no nível anterior acertam o item; e
- o ajuste da curva CCI é bom.

Em geral, os itens são agrupados por intervalos de um ou de meio desvio-padrão. Essa escolha vai depender da quantidade de itens disponíveis para análise. O ponto 500, por exemplo, representará o intervalo (450; 550), se considerarmos 1 desvio-padrão de 100 pontos. Assim, poderemos ter os níveis 1 (abaixo de 250], 2 (250;350], 3 (350;450], 4 (450;550], 5 (550;650], 6 (650; 750], 7 (750, 850], 8 (acima de 850).

Os especialistas das áreas avaliadas que farão a interpretação da escala recebem os itens separados por níveis, cada item com seu enunciado completo, suas estatísticas clássicas e um gráfico com sua curva característica proveniente da calibração da TRI e o gráfico da distribuição por alternativas da proporção de acertos por faixa da escala. De posse dessas informações, são feitas as descrições dos níveis de acordo com os exemplos apresentados no capítulo 1 destas notas. Fica, então, estabelecido o que os indivíduos demonstraram ter desenvolvido em termos das habilidades avaliadas de acordo com o nível em que eles se encontram.

Alguns problemas podem ocorrer nesse processo como quantidade insuficiente de itens para caracterizar/interpretar cada ponto da escala. Além disso, os níveis extremos da escala, referentes às habilidades mais baixas e às mais altas, são, de modo geral, mal caracterizados, por serem definidos, respectivamente, por itens muito fáceis ou muito difíceis, que, em geral, são poucos. Convém observar que, nos níveis extremos superiores da escala, há poucos estudantes, isto é, é possível interpretar pedagogicamente um nível superior da escala, mas uma porcentagem muito baixa dos alunos avaliados domina os conhecimentos descritos nesse nível.

Apesar disso, a interpretação da escala pode ser considerada uma “pitada” da avaliação formativa que se introduz em um processo de avaliação de larga escala, de extrema relevância para todos os participantes, sejam eles estudantes ou gestores.

Capítulo 3

Estimação dos parâmetros dos itens e das proficiências

As escalas de proficiências discutidas nas seções anteriores são obtidas a partir da calibração dos itens das provas por intermédio do modelo da TRI. Calibrar um item de um teste significa identificar seus parâmetros (discriminação, dificuldade e acerto ao acaso), o que é feito por meio de técnicas estatísticas de estimação. Os dados que temos após a aplicação de um teste a um grupo de indivíduos são as respostas que esses indivíduos dão aos itens do teste. Disso resulta uma matriz cujo número de linhas corresponde à quantidade de respondentes e o número de colunas é igual ao número de itens do teste. Se for um teste de aptidão, então as respostas serão todas iguais a 0 ou 1, significando erro ou acerto a cada item. Como conhecemos apenas as respostas dos indivíduos, podemos pensar em estimar conjuntamente os parâmetros dos itens e as habilidades dos respondentes, ou supor conhecidas as habilidades usando, por exemplo, o escore total no teste, para, em seguida, estimar os parâmetros dos itens.

3.1 Método de estimação dos parâmetros dos itens

Para fins de ilustrar o procedimento básico da TRI para resolver esse problema, trataremos nesta seção da estimação dos parâmetros dos itens quando as habilidades são conhecidas, utilizando o método da máxima verossimilhança (*maximum likelihood*), já que as probabilidades serão calculadas a partir dos dados já coletados, empíricos, ou seja, *a posteriori*. Em seguida, estimaremos as habilidades supondo conhecidos os parâmetros dos itens. Os resultados aqui apresentados foram adaptados de Andrade *et al* (2000).

O método da máxima verossimilhança é um processo que busca descobrir o valor de θ que maximiza a verossimilhança de ocorrer um padrão de respostas aos itens, obtido por parte de determinado indivíduo. Isto é, dado um padrão de respostas, por exemplo, 111100001100111, determinar a posição mais verossímil que o sujeito ocupa no contínuo dos θ s, isto é, qual o nível de aptidão mais provável do sujeito que obteve tal padrão de respostas.

Para desenvolver todo o procedimento, são feitas duas hipóteses fundamentais:

- a) independência das respostas oriundas de indivíduos diferentes;
- b) independência local, isto é, os itens são respondidos de forma independente por cada indivíduo, fixada sua habilidade.

A segunda suposição significa que, para cada valor da habilidade θ , se tomarmos um conjunto de indivíduos com essa habilidade, as covariâncias entre as respostas para cada par de itens serão nulas. (ver Lord, 1980 e Lord & Novick, 1968).

Dito de outro modo, para examinandos com uma dada aptidão, a probabilidade de obter um padrão de respostas a um conjunto de itens é igual ao produto das probabilidades das respostas de cada item. Embora pareça improvável que as respostas de um mesmo indivíduo não estejam correlacionadas, a independência local afirma que, se houver correlação, esta será devida à influência de fatores diferentes da aptidão (θ) que está sendo medida. Se esses outros fatores forem controlados, isso significa que o indivíduo responde exclusivamente em função do tamanho de seu θ , já que a única causa de resposta do sujeito é esse suposto θ dominante. Como a unidimensionalidade do teste - homogeneidade do conjunto de itens que, supostamente, devem estar medindo um único traço latente - implica a independência local (ver Hambleton & Swaminathan, 1991), se os itens forem elaborados de modo a satisfazer a suposição de unidimensionalidade, então a segunda hipótese anteriormente estabelecida é automaticamente satisfeita.

Logo, assumindo as hipóteses acima como verdadeiras, a probabilidade de um conjunto de n indivíduos j dar uma resposta $X_j = (x_{j1}, x_{j2}, \dots, x_{jl})$ a um conjunto l de itens, levando-se em consideração a aptidão θ_j de cada um e os parâmetros dos itens $\lambda_i = a_i, b_i$ ou c_i , é dada por:

$$\begin{aligned}
 L(\lambda) &= \prod_{j=1}^n \prod_{i=1}^I P(X_{ji} = x_{ji} | \theta_j, \lambda_i) \\
 &= \prod_{j=1}^n \prod_{i=1}^I P(X_{ji} = 1 | \theta_j, \lambda_i)^{x_{ji}} P(X_{ji} = 0 | \theta_j, \lambda_i)^{1-x_{ji}} \\
 &= \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{x_{ji}} Q_{ji}^{1-x_{ji}}, \tag{1}
 \end{aligned}$$

em que $P_{ji} = P(X_{ji} = 1 | \theta_j, \lambda_i)$ e $Q_{ji} = 1 - P_{ji}$ é a probabilidade de o indivíduo j errar o item i .

Como precisaremos calcular os valores λ_i que maximizam a função de verossimilhança (1), para facilitar os cálculos, trabalharemos com a log-verossimilhança, que pode ser escrita como

$$\log L(\lambda) = \sum_{j=1}^n \sum_{i=1}^I \{x_{ji} \log P_{ji} + (1 - x_{ji}) \log Q_{ji}\}. \tag{2}$$

Os estimadores de máxima verossimilhança de $\lambda_i, i = 1, \dots, I$, com $I = 3l$, são os valores que maximizam a verossimilhança, ou equivalentemente, são as soluções da equação

$$\frac{\partial \log L}{\partial \lambda_i}(\lambda) = 0, \quad i = 1, \dots, I.$$

Notemos que

$$\begin{aligned}
\frac{\partial \log L}{\partial \lambda_i} &= \sum_{j=1}^n \left\{ x_{ji} \frac{\partial (\log P_{ji})}{\partial \lambda_i} + (1 - x_{ji}) \frac{\partial (\log Q_{ji})}{\partial \lambda_i} \right\} \\
&= \sum_{j=1}^n \left\{ x_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) - (1 - x_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) \right\} \\
&= \sum_{j=1}^n \left\{ x_{ji} \frac{1}{P_{ji}} - (1 - x_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) \\
&= \sum_{j=1}^n \left\{ \frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right).
\end{aligned}$$

Usando a notação

$$W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}},$$

em que

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

obtemos, para cada um dos parâmetros do item i , as derivadas parciais de primeira ordem da função log-verossimilhança dadas por:

$$\begin{aligned}
\frac{\partial \log L}{\partial a_i} &= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial a_i} \right) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) D(1 - c_i)(\theta_j - b_i) P_{ji}^* Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= D(1 - c_i) \sum_{j=1}^n (x_{ji} - P_{ji})(\theta_j - b_i) W_{ji}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \log L}{\partial a_i} &= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial a_i} \right) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji})(-1) Da_i(1 - c_i) P_{ji}^* Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= -Da_i(1 - c_i) \sum_{j=1}^n (x_{ji} - P_{ji}) W_{ji}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \log L}{\partial c_i} &= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) \left(\frac{\partial P_{ji}}{\partial c_i} \right) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) Q_{ji}^* \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \\
&= \sum_{j=1}^n \left\{ (x_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} \right\}.
\end{aligned}$$

Assim, as equações de estimação para os parâmetros a_i , b_i , c_i são, respectivamente,

$$\sum_{j=1}^n (x_{ji} - P_{ji})(\theta_j - b_i)W_{ji} = 0, \quad (3)$$

$$\sum_{j=1}^n (x_{ji} - P_{ji})W_{ji} = 0, \quad (4)$$

$$\sum_{j=1}^n (x_{ji} - P_{ji})\frac{W_{ji}}{P_{ji}^*} = 0. \quad (5)$$

Na maioria das vezes, as condições obtidas das derivadas de primeira ordem para o problema de maximização da função de verossimilhança não permitem a obtenção de uma solução explícita para os estimadores em questão. O sistema de equações gerado é quase sempre não linear, obrigando que a maximização seja feita por algum processo numérico. Os procedimentos de otimização numéricos funcionam de forma recursiva, sendo os valores dos parâmetros na etapa $k + 1$ uma função de seus valores na etapa k . O algoritmo numérico consiste em tentar um valor para cada parâmetro e, depois, corrigi-lo sucessivamente até que algum critério de convergência seja atendido, quando, então, chega-se a um máximo para a função de verossimilhança. Em alguns casos, quando não ocorre convergência, o processo de iteração tem de ser interrompido depois de um número de iterações preestabelecido. Isso é exatamente o que ocorre com as equações (3), (4) e (5), que não possuem solução explícita. Podemos, por exemplo, utilizar o método de Newton-Raphson para fazer as estimações desejadas. Este é o tema da seção seguir.

3.2 Aplicação do método de Newton-Raphson

O método de Newton-Raphson foi introduzido para resolver equações do tipo $f(x) = 0$, encontrando, por meio de sucessivas iterações, valores que se aproximam cada vez mais da solução. De modo geral,

$$x_{k+1} = x_k - \frac{f(x_{k+1})}{f'(x_k)},$$

considerando-se a etapa k e a aproximação seguinte. Esse processo iterativo é continuado até que o valor absoluto da diferença $x_{k+1} - x_k$ seja pequeno o suficiente ou até que seja feito um número máximo de iterações preestabelecido.

No nosso caso, a função $l(\lambda) = \log L(\lambda)$ possui várias variáveis, λ_i , com $i = 1 \dots 3l$, e precisaremos encontrar os valores dos parâmetros que anulam as primeiras derivadas de $l(\lambda)$, o que corresponde a obter as soluções das equações de estimação dos parâmetros. Tomando valores iniciais $\hat{\lambda}_i^0 = (a_i^0, b_i^0, c_i^0)'$, a primeira estimativa pode ser escrita como $\hat{\lambda}_i^1 = \hat{\lambda}_i^0 + \Delta \hat{\lambda}_i^0$, isto é,

$$\begin{aligned} \hat{a}_i^1 &= \hat{a}_i^0 + \Delta \hat{a}_i^0, \\ \hat{b}_i^1 &= \hat{b}_i^0 + \Delta \hat{b}_i^0, \\ \hat{c}_i^1 &= \hat{c}_i^0 + \Delta \hat{c}_i^0, \end{aligned} \quad (6)$$

em que $\Delta\hat{a}_i^0, \Delta\hat{b}_i^0$, e $\Delta\hat{c}_i^0$ são erros de aproximação. Fazendo a expansão em série de Taylor de $\partial l(\lambda)/\partial\lambda_i$ em uma vizinhança de $\hat{\lambda}_i^0$, obtemos

$$\begin{aligned}\frac{\partial l}{\partial a_i}(\lambda) &= \frac{\partial l}{\partial a_i}(\hat{\lambda}_i^0) + \Delta\hat{a}_i^0 \frac{\partial^2 l}{\partial a_i^2}(\hat{\lambda}_i^0) + \Delta\hat{b}_i^0 \frac{\partial^2 l}{\partial a_i \partial b_i}(\hat{\lambda}_i^0) + \Delta\hat{c}_i^0 \frac{\partial^2 l}{\partial a_i \partial c_i}(\hat{\lambda}_i^0) + R_{a_i} \\ \frac{\partial l}{\partial b_i}(\lambda) &= \frac{\partial l}{\partial b_i}(\hat{\lambda}_i^0) + \Delta\hat{b}_i^0 \frac{\partial^2 l}{\partial b_i^2}(\hat{\lambda}_i^0) + \Delta\hat{b}_i^0 \frac{\partial^2 l}{\partial b_i \partial a_i}(\hat{\lambda}_i^0) + \Delta\hat{c}_i^0 \frac{\partial^2 l}{\partial b_i \partial c_i}(\hat{\lambda}_i^0) + R_{b_i} \\ \frac{\partial l}{\partial c_i}(\lambda) &= \frac{\partial l}{\partial c_i}(\hat{\lambda}_i^0) + \Delta\hat{a}_i^0 \frac{\partial^2 l}{\partial c_i^2}(\hat{\lambda}_i^0) + \Delta\hat{b}_i^0 \frac{\partial^2 l}{\partial c_i \partial a_i}(\hat{\lambda}_i^0) + \Delta\hat{c}_i^0 \frac{\partial^2 l}{\partial c_i \partial b_i}(\hat{\lambda}_i^0) + R_{c_i},\end{aligned}$$

Nessas expressões, estamos usando que $\partial l/\partial\lambda_i$ é uma função apenas de λ_i , não dependendo de λ_l para $l \neq i$.

Vamos agora introduzir a seguinte notação

$$\begin{aligned}L_1 &= \frac{\partial l}{\partial a_i} & L_{11} &= \frac{\partial^2 l}{\partial a_i^2} & L_{12} &= \frac{\partial^2 l}{\partial a_i \partial b_i} & L_{13} &= \frac{\partial^2 l}{\partial a_i \partial c_i}, \\ L_2 &= \frac{\partial l}{\partial b_i} & L_{21} &= \frac{\partial^2 l}{\partial b_i \partial a_i} & L_{22} &= \frac{\partial^2 l}{\partial b_i^2} & L_{23} &= \frac{\partial^2 l}{\partial b_i \partial c_i}, \\ L_3 &= \frac{\partial l}{\partial c_i} & L_{31} &= \frac{\partial^2 l}{\partial c_i \partial a_i} & L_{32} &= \frac{\partial^2 l}{\partial c_i \partial b_i} & L_{33} &= \frac{\partial^2 l}{\partial c_i^2},\end{aligned}$$

em que todas as funções são avaliadas no pontos $(\hat{\lambda}_i^0)$. Desprezando os restos no desenvolvimento de Taylor, obtemos

$$\begin{aligned}0 &= L_1 + L_{11}\Delta\hat{a}_i^0 + L_{12}\Delta\hat{b}_i^0 + L_{13}\Delta\hat{c}_i^0, \\ 0 &= L_2 + L_{12}\Delta\hat{a}_i^0 + L_{22}\Delta\hat{b}_i^0 + L_{23}\Delta\hat{c}_i^0, \\ 0 &= L_3 + L_{13}\Delta\hat{a}_i^0 + L_{23}\Delta\hat{b}_i^0 + L_{33}\Delta\hat{c}_i^0.\end{aligned}$$

O sistema acima pode ser escrito, em notação matricial, do seguinte modo:

$$-\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} \Delta\hat{a}_i^0 \\ \Delta\hat{b}_i^0 \\ \Delta\hat{c}_i^0 \end{pmatrix}.$$

Assim, a primeira estimação para os parâmetros pode ser obtida pela seguinte equação

$$\begin{pmatrix} \hat{a}_i^1 \\ \hat{b}_i^1 \\ \hat{c}_i^1 \end{pmatrix} = \begin{pmatrix} \hat{a}_i^0 \\ \hat{b}_i^0 \\ \hat{c}_i^0 \end{pmatrix} - \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}^{-1} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}.$$

Como o processo é iterativo, após obtido $\hat{\lambda}_i^1$, considera-se este o novo valor inicial para a obtenção de $\hat{\lambda}_i^2$, repetindo-se o processo até que algum critério de parada seja alcançado.

Para explicitar os valores de L_{kl} , com $k, l = 1, 2, 3$, precisamos calcular as derivadas de segunda ordem de $l = \log L$ em relação a λ_i . Fazendo isso, obtemos

$$\frac{\partial \log L}{\partial \lambda_i \lambda_i'} = \sum_{j=1}^n \left\{ \left[\frac{\partial}{\partial \lambda_i} \left(\frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \right] \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right)' + \left(\frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) \left(\frac{\partial^2 P_{ji}}{\partial \lambda_i \partial \lambda_i'} \right) \right\}.$$

Calculando a derivada do quociente explicitado no primeiro termo do somatório acima, encontramos

$$\begin{aligned}
\frac{\partial}{\partial \lambda_i} \left(\frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) &= \\
&= \frac{1}{(P_{ji} Q_{ji})^2} \left\{ -P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) - (x_{ji} - P_{ji}) \left(\frac{\partial P_{ji} Q_{ji}}{\partial \lambda_i} \right) \right\} \\
&= \frac{-1}{(P_{ji} Q_{ji})^2} \left\{ P_{ji} Q_{ji} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) + (x_{ji} - P_{ji}) \left[\left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) - 2P_{ji} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) \right] \right\} \\
&= \frac{-1}{(P_{ji} Q_{ji})^2} \{ P_{ji} Q_{ji} + (x_{ji} - P_{ji})(1 - 2P_{ji}) \} \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right) \\
&= \frac{-1}{(P_{ji} Q_{ji})^2} (x_{ji} - P_{ji})^2 \left(\frac{\partial P_{ji}}{\partial \lambda_i} \right).
\end{aligned}$$

A última igualdade segue do fato que $x_{ji} = x_{ji}^2$, já que x_{ji} assume somente os valores 0 e 1. Substituindo esse resultado na equação anterior, obtemos o resultado desejado. Fazendo iterações sucessivas, chegamos aos valores de estimação com a aproximação que for pré-fixada.

3.3 Estimação das habilidades

Nesta seção, vamos tratar da estimação das habilidades quando os parâmetros dos itens são conhecidos. Os itens já podem ter vindo de algum processo de pré-testagem ou optou-se mesmo pela estimação em separado.

Usando novamente as hipóteses da independência entre as respostas de diferentes indivíduos e da independência local, podemos escrever a log-verossimilhança como em (2), agora como função de θ e não de λ , ou seja,

$$\log L(\theta) = \sum_{j=1}^n \sum_{i=1}^I \{ x_{ji} \log P_{ji} + (1 - x_{ji}) \log Q_{ji} \}. \quad (7)$$

O estimador de máxima verossimilhança de θ_j , para $j = 1, \dots, n$, é o valor que anula a derivada primeira de $\log L$ em relação a θ_j , calculada a seguir.

$$\begin{aligned}
\frac{\partial \log L}{\partial \theta_j} &= \sum_{i=1}^I \left\{ x_{ji} \frac{\partial(\log P_{ji})}{\partial \theta_j} + (1 - x_{ji}) \frac{\partial(\log Q_{ji})}{\partial \theta_j} \right\} \\
&= \sum_{i=1}^I \left\{ x_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) - (1 - x_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial Q_{ji}}{\partial \theta_j} \right) \right\} \\
&= \sum_{i=1}^I \left\{ x_{ji} \frac{1}{P_{ji}} - (1 - x_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) \\
&= \sum_{i=1}^I \left\{ \frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) \\
&= \sum_{i=1}^I \left\{ (x_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right),
\end{aligned}$$

Como

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

obtemos

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

Segue, então, que a equação de estimação para θ_j , $j = 1, \dots, n$, é igual a

$$D \sum_{i=1}^I a_i (1 - c_i) (x_{ji} - P_{ji}) W_{ji} = 0. \quad (8)$$

Assim como nas equações (3), (4) e (5), a equação (8) também não apresenta solução explícita para θ_j e, por isso, precisamos de algum método iterativo para obter as estimativas desejadas. Com raciocínio análogo ao feito com os parâmetros dos itens e considerando $\hat{\theta}_j^k$ a estimativa de θ_j na iteração k do método de Newton-Raphson, obtemos

$$\hat{\theta}_j^{k+1} = \hat{\theta}_j^k - \left[\frac{\partial^2 \log L}{\partial \theta_j^2} (\hat{\theta}_j^k) \right]^{-1} \frac{\partial \log L}{\partial \theta_j} (\hat{\theta}_j^k).$$

Para calcular os valores de $\hat{\theta}_j^{k+1}$, usamos as seguintes expressões das derivadas de primeira e segunda ordem da função $\log L$:

$$\begin{aligned}
\frac{\partial \log L}{\partial \theta_j} &= \sum_{i=1}^I \left\{ (x_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) h_{ji} \\
&= \sum_{i=1}^I (x_{ji} - P_{ji}) W_{ji} h_{ji}
\end{aligned}$$

e

$$\begin{aligned}\frac{\partial^2 \log L}{\partial \theta_i^2} &= \sum_{i=1}^I \left\{ \left(\frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) (P_{ji}^* Q_{ji}^*) H_{ji} - \left(\frac{x_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 (P_{ji}^* Q_{ji}^*)^2 h_{ji}^2 \right\} \\ &= \sum_{i=1}^I (x_{ji} - P_{ji}) W_{ji} \{ H_{ji} - (x_{ji} - P_{ji}) W_{ji} h_{ji}^2 \}\end{aligned}$$

com

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) = D a_i (1 - c_i)$$

e

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) = D^2 a_i^2 (1 - c_i) (1 - 2P_{ji}^*).$$

Esses cálculos ilustram a metodologia subjacente à teoria de resposta ao item. Há muitas situações a serem consideradas, não somente a respeito da estimação de parâmetros e habilidades mas também sobre erro padrão das medidas, problemas de convergência dos métodos, aplicações de testes a mais de um grupo de indivíduos, aplicações de testes a grupos diferentes mas com itens-âncora, equalização dos resultados, viés de itens, multidimensionalidade dos testes, fidedignidade, validade, viés, modelos para itens não dicotômicos etc. O leitor que desejar explorar mais detalhes matemáticos relativos à teoria encontrará em Andrade *et al* (2000) uma excelente fonte de leitura.

Capítulo 4

Considerações Finais

A realidade brasileira mostra milhões de crianças e jovens que, ano após ano, experimentam o desânimo, a frustração e o abandono escolar. Isso sugere que, subjacente à avaliação, há uma teia de significados que merecem ser considerados em dimensões amplas e diversificadas e que refletem a complexidade da própria subjetividade dos atores envolvidos: quem avalia e quem é avaliado. O processo não é neutro e sofre influência das relações sociais estabelecidas dentro e fora do contexto escolar.

A análise feita nestas notas mostra que o Brasil já acumula uma experiência considerável com relação aos processos de avaliação de larga escala e revela a complexidade das diversas etapas dos processos, que engloba a qualidade dos instrumentos utilizados, as estratégias e situações pelas quais se obtêm informações acerca do aprendizado dos estudantes e os métodos utilizados para interpretar o desempenho. O INEP, desde a criação do SAEB, em 1990, vem produzindo indicadores sobre o sistema educacional brasileiro. Alguns deles apontam para problemas graves na eficiência do ensino oferecido pelas redes de escolas brasileiras, como os baixos desempenhos em leitura e na apreensão de conceitos básicos de matemática, por exemplo. Em face de tal realidade, o governo federal e muitos governos estaduais e municipais têm empreendido esforços para reverter esse quadro. Nesse sentido, a opção pelo uso da TRI para a construção de bancos de itens calibrados e de escalas de proficiência comparáveis ao longo do tempo proporciona inferências relevantes com relação às aprendizagens dos estudantes e à análise da eficácia das ações de melhoria da qualidade de ensino implantadas.

A solidez com que o SAEB foi construído influenciou a escolha de metodologias similares para os demais processos – ENCCEJA, ENEM, Prova Brasil, Provinha Brasil e PNCICD –, abrindo também para estes as possibilidades futuras de acompanhamento longitudinal dos desempenhos e interpretação das escalas de proficiência. Já na educação superior, a diversidade de cursos e instituições, aliada à explosão da oferta do ensino a distância, trazem para o SINAES desafios diferenciados daqueles encontrados

na avaliação da educação básica e impõem a necessidade de aprofundamento nos modelos de instrumentos utilizados e nas estratégias de coleta de dados.

Apesar dos avanços alcançados ao longo desses anos, muitos obstáculos ainda precisam ser transpostos, principalmente em um país de dimensões continentais e que possui a quinta maior população mundial, com desigualdades profundas no que se referem à formação e acesso a cultura, educação e uso de novas tecnologias. Isso coloca a necessidade de cada vez mais se fazer uma avaliação consistente com os desafios curriculares da educação contemporânea e de se criarem metodologias de ensino que efetivamente contemplem alguma possibilidade de mudança qualitativa.

Desse modo, torna-se necessário: aprofundar o caráter de diagnóstico das avaliações de larga escala; incluir itens abertos, de resposta construída pelos estudantes, que avaliam competências distintas das associadas a itens objetivos; aprimorar critérios de correção/codificação de itens abertos; criar testes adaptativos, com possibilidade de inserção de áudio, vídeo e itens abertos de resposta curta, o que reduzirá, também, os custos com logística de aplicação; usar a TRI para uniformizar melhor as correções de itens discursivos; otimizar o processo de construção dos bancos de itens, mantendo-se o sigilo inerente ao processo; expandir o uso de itens abertos para outros processos de avaliação, com o uso de “metodologia de correção *online* descentralizada e acompanhamento em tempo real” (Andrade e Rabelo, 2006); simplificar a logística de aplicação com o uso de tecnologias adequadas; reduzir o tempo gasto com análise do desempenho para agilizar o processo de implantação de políticas públicas de melhoria da qualidade do ensino; atualizar as matrizes de referência (SAEB, por exemplo), sem perder a comparabilidade nem a interpretação das escalas de proficiência; aperfeiçoar os mecanismos de adaptação dos testes para deficientes visuais e auditivos, em uma política inclusiva; repensar as devolutivas para estudantes, professores e gestores, com maior clareza e simplicidade de informação, de modo a contribuírem para regular o processo ensino-aprendizagem.

Enfim, fazer com que o processo se aproxime cada vez mais daquilo que os educadores esperam da avaliação, mesmo que seja de larga escala: que ela esteja a serviço da aprendizagem.

Referências Bibliográficas

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

ANDRADE, G.G. & RABELO, M.L. *Os desafios da avaliação de textos em larga escala*. In: III Congresso Internacional de Avaliação Educacional. Eixo Temático 3 – Avaliação de Ensino-Aprendizagem no Ensino Médio. Fortaleza, 76-92, 2006.

_____. *O ENEM e os desafios da correção das redações*. In: ANDRADE, G.G. & RABELO, M.L. (org.) *A Produção de Textos no ENEM: desafios e conquistas*. Brasília: EdUnB. (Desafios e Conquistas, vol. 1), 2007.

BAKER, F.B. *Item Response Theory: Parameter Estimation Techniques*. Marcel Dekker, 1992.

BEATON, A.E., ALLEN, N.L. *Interpreting scales through scale anchoring*. Journal of Educational Statistics, v.17, p.191-204, 1992.

BRASIL. *Constituição da República Federativa do Brasil de 1988* (Atualizada). Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L8069.htm. Acesso em: 01 mar. 2010.

_____. *Encceja - Exame Nacional para Certificação de Competências de Jovens e Adultos*. Coleção de Livros do Professor e do Estudante: Áreas de Linguagens e Códigos, Matemática, Ciências Humanas e Ciências da Natureza. Brasília: Inep/MEC, 2003. 12 v.

_____. *Encceja - Exame Nacional para Certificação de Competências de Jovens e Adultos*. Coleção de Livros do Estudante: Áreas de Linguagens e Códigos, Matemática, Ciências Humanas e Ciências da Natureza. Brasília: Inep/MEC. 2. ed. Brasília: Inep/MEC, 2006, 8 v.

_____. FERNANDES, R. *Índice de Desenvolvimento da Educação Básica (IDEB): metas intermediárias para a sua trajetória no Brasil, estados, municípios e escolas*. Nota Técnica. Disponível em: <<http://www.inep.gov.br>>. Acesso em: 15 set. 2009.

_____. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Matriz de Referência para a Prova Nacional de Concurso para o Ingresso na Carreira Docente*. Disponível em: http://download.inep.gov.br/download/basica/concurso_docente/proposta_matriz_discussao_ versao_atualizada_13042011.pdf. Acesso em: 05 jul. 2011.

_____. Lei nº 9.394, de 20 de dezembro de 1996. *Lei de Diretrizes e Bases da Educação Nacional* (Atualizada). Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm. Acesso em: 01 mar. 2010.

_____. Lei nº 10.861, de 14 de abril de 2004. *Institui o Sistema Nacional de Avaliação da Educação Superior – SINAES*. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/110.861.htm. Acesso em: 05 jun. 2011.

_____. Ministério da Educação. *ENEM - Exame Nacional do Ensino Médio*. Documento Básico, 2000. Brasília: Inep/MEC, 1999.

_____. Ministério da Educação. *Eixos Cognitivos do Enem*. Brasília: Inep/MEC, 2007(reimpressão).

_____. Ministério da Educação. *Médias de desempenho do SAEB/2005 em perspectiva comparada*. Brasília: Inep/MEC, 2006.

_____. Ministério da Educação. *Plano Decenal de Educação para Todos*, Brasília: MEC, 1993.

_____. Portaria Inep/MEC n.º 109, de 27 de maio de 2009, publicada no D.O.U. de 8 de junho de 2009, Seção 1, 14-20.

_____. Portaria Inep/MEC n.º 174, de 31 de julho de 2009, publicada no D.O.U. de 5 de agosto de 2009, Seção 1, 10-13.

_____. Portaria n.º 2.270, de 14 de agosto de 2002. *Institui o Exame Nacional de Certificação de Competências de Jovens e Adultos (Encceja)*. Disponível em: <http://portal.inep.gov.br/web/encceja/legislacao>. Acesso em: 05 jun. 2011.

_____. Portaria n.º 2.051, de 9 de julho de 2004. *Regulamenta os procedimentos de avaliação do Sistema Nacional de Avaliação da Educação Superior (SINAES)*, Disponível em: www.inep.gov.br/superior/enade/. Acesso em: 05 jun. 2011.

_____. *Médias de desempenho do SAEB/2005 em perspectiva comparada*. Brasília: Inep/MEC, 2006.

_____. *Provinha Brasil*. Disponível em: <http://portal.inep.gov.br/web/provinha-brasil/caracteristicas>. Acesso em: 05 jul. 2011.

HAMBLETON, R.K., SWAMINATHAN, H. & ROGERS, H.J. *Fundamentals of Item Response Theory*. Sage Publications, 1991.

LE BOTERF, G *Desenvolvendo a competência dos profissionais*. Porto Alegre: Artmed, 2003.

MARINHO-ARAÚJO, C.M. *O Desenvolvimento de Competências no ENADE: a mediação da avaliação nos processos de desenvolvimento psicológico e profissional*. Avaliação. Revista da Rede de Avaliação Institucional da Educação Superior - RAIES. Campinas: Unicamp, v.9, n.4, 77-97, 2004.

MARINHO-ARAÚJO, C.M., POLIDORI, M.M., FONSECA, D.G., OLIVEIRA, C.B., OLIVEIRA, G., SOARES, P.G., CURÇO, S., RODRIGUES, C. & REIS, J. *Processos avaliativos e o desenvolvimento de competências*. Actas do VII Simpósio Nacional de Investigação em Psicologia da Universidade do Minho, Portugal, 2947-2964, 2010.

PASQUALI, L. *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Vozes, 2003.

RABELO, M.L. *A matriz de matemática do SAEB e o contexto escolar*. Pesquisa & Avaliação: Revista da Escola SAERJ 2009, 1, 20-27, 2010.

RABELO, M. L. *A Teoria de Resposta ao Item no Novo Enem*. Explicando o Enem - Educar para as Competências. São Paulo: Abril Educação, 65-67, 2009.

RABELO, M.L. *Matriz de referência como centro dos processos de avaliação em larga escala*. Pesquisa & Avaliação: Revista do Professor: SADEAM 2010: Matemática e suas tecnologias, 2, 2011.

RABELO, M.L. *Os Exames Nacionais de Avaliação e a Teoria de Resposta ao Item*. Minicurso apresentado no I Colóquio de Matemática da Região do Centro-Oeste. Campo Grande - MS, 2009.

RABELO, M.L. *Avaliação de larga escala em perspectiva comparada: a experiência brasileira*. VIII Congresso Iberoamericano de Avaliação. Portugal, 2011.

ZARIFIAN, P. *O modelo da competência: trajetória histórica, desafios atuais e propostas*. São Paulo: Editora Senac, 2003.

