



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Instituto de Física
Programa de Pós-Graduação em Ensino de Física
Mestrado Profissional em Ensino de Física

AVALIAÇÕES EM LARGA ESCALA E O PROFESSOR DE FÍSICA

Wanderley Paulo Gonçalves Junior

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ensino de Física, Instituto de Física, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Ensino de Física.

Orientadora: Marta Feijó Barroso

Rio de Janeiro
02/2012

AVALIAÇÕES DE LARGA ESCALA E O PROFESSOR DE FÍSICA

Wanderley Paulo Gonçalves Junior

Orientadora:
Marta Feijó Barroso

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Ensino de Física, Instituto de Física, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Ensino de Física.

Aprovada por:

Dr. Marta Feijó Barroso (Presidente)

Dr. Alicia Maria Catalano de Bonamino (PUC-Rio)

Dr. Samuel Jurkiewicz (COPPE-UFRJ)

Dr. Antonio Carlos Fontes dos Santos (IF-UFRJ)

Rio de Janeiro
02/2012

FICHA CATALOGRÁFICA

G635a Gonçalves Junior, Wanderley Paulo
Avaliação de Larga Escala e o Professor de Física /
Wanderley Paulo Gonçalves Junior - Rio de Janeiro: UFRJ / IF, 2012.
xi, 227 f.: il.;30cm.
Orientador: Marta Feijó Barroso
Dissertação (mestrado) – UFRJ / Instituto de Física /
Programa de Pós-Graduação em Ensino de Física, 2012.
Referências Bibliográficas: f. 121-125.
1. Ensino de Física. 2. Avaliação de Larga Escala. 3. Teoria
de Resposta ao Item. I. Barroso, Marta Feijó. II. Universidade Federal
do Rio de Janeiro, Instituto de Física, Programa de Pós-Graduação
em Ensino de Física. III. Explorando o Mundo das Avaliações de
Larga Escala.

Dedicatória

*Aos meus pais, Wanderley e Alzira,
e a todos que estiveram comigo
em mais essa conquista.*

Agradecimentos

À minha orientadora, Marta, a quem devo minha formação como pesquisador. Pela paciência, dedicação, exigência e comprometimento com o trabalho.

Aos meus familiares, Cristian, Carmem, Simone, Edilaine, Samantha, Cassiano, Raphael, Gabriel, Talitha, Ariane, Gabriela, Helena, Cristiane, Sabrina e Alan. Pelo constante apoio e suporte afetivo.

Aos colegas do grupo de pesquisa: Marcelo, Gustavo e Luciana. As discussões, críticas e almoços fizeram parte da história dessa pesquisa.

Aos meus colegas e amigos de mestrado: Almir, Felipe, José Fernando, José Luiz, Leonardo, Magali, Otoniel, Paulo Victor, Reynaldo, Sandro e Vitor. Os debates em sala de aula, os esforços em conjunto para cumprirmos as tarefas, o companheirismo, a amizade, o prazer de ter convivido com vocês, tudo levo comigo em meu peito.

Enfim, a todos os amigos, colegas e mestres que estiveram presentes neste percurso...

RESUMO

AVALIAÇÕES EM LARGA ESCALA E O PROFESSOR DE FÍSICA

Wanderley Paulo Gonçalves Junior

Orientadora:
Marta Feijó Barroso

Resumo da Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Ensino de Física, Instituto de Física, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Ensino de Física.

Os processos avaliativos constituem-se peças fundamentais no quebra-cabeça do ensino e aprendizagem de ciências. Este trabalho se propõe tentar promover, nos professores, uma reflexão em relação a sistemas de avaliação, em especial os de larga escala e, através de informações e exemplos que buscam facilitar sua compreensão, tornando-os agentes desse novo processo que os envolve. No Capítulo 2, faz-se, inicialmente, uma breve apresentação sobre as funções das avaliações, discorrendo-se em seguida sobre algumas avaliações de larga escala internacionais e nacionais. No Capítulo 3 se discute a possibilidade de se realizar medidas nas áreas de ciências sociais, focando a mensuração de variáveis não observáveis, finalizando com uma apresentação sobre a Teoria Clássica dos Testes e a Teoria de Resposta ao Item. Enfim nos Capítulos 4 e 5 encontra-se o coração do trabalho, ou seja, os resultados dos trabalhos desenvolvidos na aplicação de um questionário de física térmica e na análise de provas de 2009 a 2011 do Exame nacional do Ensino Médio.

Palavras-chave: Ensino de Física, Avaliação de Larga Escala, Teoria de Resposta ao Item

Rio de Janeiro
02/2012

ABSTRACT

LARGE-SCALE ASSESSMENT AND PHYSICS TEACHER

Wanderley Paulo Gonçalves Junior

Supervisor:
Marta Feijó Barroso

Abstract of master's thesis submitted to Programa de Pós-Graduação em Ensino de Física, Instituto de Física, Universidade Federal do Rio de Janeiro, in partial fulfillment of the requirements for the degree Mestre em Ensino de Física.

Learning assessments are fundamental to teachers. In this work, we propose a reflexion on assessment systems and processes, particularly large scale assessments, in order to provide teachers information that allow them to understand and use the results of assessment results to improve their professional activity. We begin in Chapter 2 with a review on the roles of assessments, and of large scale assessments in Brazil and worldwide. In Chapter 3, we discuss processes of measurement in social sciences, and present a discussion on the Classical Test Theory and on Item Response Theory as possible methodologies of providing measurements of student's learning. Finally we present the results obtained in two subjects: the use of a thermal physics questionnaire in high school, and the analysis of the ENEM, the national test in Brazil at the end of high school.

Keywords: Physics Teaching, Large-Scale Assessment, Item Response Theory

Rio de Janeiro
02/2012

Sumário

Capítulo 1 – Introdução	1
Capítulo 2 – Avaliações de Larga Escala	4
2.1 – Aspectos Gerais sobre Avaliação de Aprendizagem.....	4
2.1.1. Avaliação Formativa	6
2.1.2. Avaliação Somativa	8
2.1.3. Avaliação Somativa com Responsabilização (“accountability”)	8
2.2 – Aspectos das Avaliações de Larga Escala	9
2.3 – Avaliações Internacionais de Larga Escala	11
2.3.1. PIRLS, TIMMS – A IEA	12
2.3.2. PIRLS – Progresso no Letramento de Leitura	14
2.3.3. TIMSS - Tendências Internacionais nos Estudos de Matemática e Ciências	15
2.3.4. PISA – Programa Internacional de Avaliação dos Estudantes	18
2.4- Avaliações de Larga Escala Nacionais	21
2.4.1. O SAEB e a Prova Brasil	21
2.4.2. O Índice de Desenvolvimento da Educação Básica (Ideb)	23
2.4.3. O Exame Nacional do Ensino Médio (ENEM)	25
2.5- Comentários finais sobre processos de Avaliação de Larga Escala.....	30
Capítulo 3 – Medidas e Avaliações de Aprendizagem	35
3.1- Medidas em Ciências Exatas e em Ciências Sociais	35
3.1.1. A Mensuração nas Ciências Sociais	36
3.1.2. A Operacionalização de Conceitos Abstratos	40

3.2- Exemplos de Medidas	43
3.2.1. Aprendizagem de Conceitos Relativos ao Movimento de Corpos Próximos à Superfície da Terra – a Aceleração da Gravidade	44
3.2.2. Mudança de Fase – Existe Água a Zero Graus?	47
3.2.3. Conclusão	49
3.3- As técnicas de avaliação de larga escala: TCT e TRI	49
3.3.1. A Teoria Clássica de Testes (TCT)	50
3.3.2. A Teoria de Resposta ao Item (TRI)	53
3.3.3. Modelo de Rasch para itens dicotômicos (Compreendendo a idéia de uma ICC)	57
3.3.4. Modelos de TRI	65
3.3.4.1. Modelo Logístico de Um Parâmetro	66
3.3.4.2. Modelo Logístico de Dois Parâmetros.....	70
3.3.4.3. Modelo Logístico de Três Parâmetros.....	71
Capítulo 4 – Questionários de Aprendizagem de Física Térmica	74
4.1- Testes Conceituais para o Ensino de Física	74
4.2- Um Teste pra Avaliar Conceitos Introdutórios de Física Térmica.....	75
4.3- Aplicação do Teste a Estudantes do Ensino Médio no Rio de Janeiro	78
4.4- Metodologia de Análise das Respostas	79
4.5- Apresentação e Análise dos Resultados	80
4.6- Considerações Finais	90
Capítulo 5 – Um Estudo do ENEM – Física	92
5.1- Descrição Metodológica	92
5.2- Resultados	95

5.2.1. O Número de Questões de Física	95
5.2.2. A Distribuição das Questões pelas Competências Previstas	96
5.2.3. As habilidades atribuídas aos itens	99
5.2.4. Os Objetos de Conhecimento	104
5.2.5. A Extensão dos Itens do ENEM	109
5.2.6. A Utilização de Imagens, Gráficos e Tabelas no Texto- Base	111
5.2.7. Nível de Contextualização dos Itens	113
5.2.8. Exigência de Conhecimentos Específicos de Física	113
5.2.9. Natureza dos Itens do ENEM	116
5.3. Considerações Finais	117
Capítulo 6 – Considerações Finais	118
Referências Bibliográficas	121
Anexo 1 – ENEM: Documento Básico (1998)	126
Anexo 2 – Matriz de Referência para o ENEM 2009	138
Apêndice 1 – Questionário de Física Térmica – Versão em Português .	164
Apêndice 2 – Relação entre Conceito e Item	168
Apêndice 3 – Ordenação dos Objetos de Conhecimento Associados a Matriz de Referência de Ciências da Natureza e suas Tecnologias	169
Apêndice 4 – Levantamento de Dados – ENEM 2009 a 2011	171
Apêndice 5 – Explorando o Mundo das Avaliações de Larga Escala.....	176

CAPÍTULO 1. INTRODUÇÃO

Os processos avaliativos constituem peças fundamentais no quebra-cabeça do ensino e aprendizagem. Apesar disso, muitos professores, ao construir suas avaliações, simplesmente reproduzem o que vivenciaram como alunos, seja na formação básica ou na universidade, conhecendo e refletindo muito pouco sobre este processo.

Enquanto estas avaliações se restringiam às paredes de sala de aula, cada professor, seguindo sua intuição, criava suas próprias interpretações dos métodos vivenciados por ele e aplicava-os, muitas vezes sem uma postura crítica e reflexiva sobre os processos avaliativos utilizados.

No entanto, a partir de 1990 com o SAEB (Sistema de Avaliação da Educação Básica) e a partir de 1998 com o ENEM (Exame Nacional do Ensino Médio), teve início uma crescente proposta de política pública de avaliação externa da aprendizagem dos alunos e da qualidade do ensino brasileiro. O estabelecimento destes exames gerou uma grande preocupação dos professores em relação às suas práticas avaliativas, uma vez que estas passaram a ser indiretamente avaliadas a nível nacional. Esta preocupação, principalmente no caso do ensino médio, intensificou-se em 2010, quando o ENEM passou a ser a principal forma de acesso à maioria das universidades públicas.

Dentro deste contexto surge a motivação para a elaboração e o desenvolvimento deste trabalho: tentar promover, entre os professores, uma reflexão em relação à avaliação, em especial aos métodos de larga escala e, através de informações e exemplos, torná-los agentes desse novo processo.

Faz-se no Capítulo 2 uma breve apresentação sobre as funções das avaliações, cujo o objetivo é localizar as avaliações de larga escala dentro dos processos avaliativos. Em seguida, faz-se um levantamento sobre algumas avaliações de larga escala internacionais em ciências e matemática e suas características (o PIRLS, o TIMMS e o PISA), percebendo que estes processos não são invenções recentes e sim frutos de pesquisas que vêm sendo desenvolvidas ao longo do tempo. Finalizando, então, este capítulo,

apresentam-se as avaliações de larga escala que vêm sendo aplicadas aos estudantes brasileiros, conhecimento tido como essencial, tendo em vista que todos os professores da educação básica possivelmente terão seus alunos avaliados por elas.

No Capítulo 3, apresenta-se uma discussão que evidencia a possibilidade de se realizar medidas nas áreas de ciências sociais, semelhante ao que é feito nas áreas de ciências exatas, desde que se leve em conta as peculiaridades da atividade científica nessas diferentes áreas. Neste contexto, discute-se a mensuração de variáveis não observáveis, como por exemplo a aprendizagem em física, através da operacionalização destes conceitos. Em seguida, discute-se rapidamente os modelos da psicometria, com a apresentação das dificuldades e limitações existentes na Teoria Clássica dos Testes (TCT), evidenciando os motivos que levaram à busca de outras teorias que permitissem a comparação de grupos quaisquer e em momentos diferentes. Apresenta-se brevemente uma teoria que se propõe a suprir as limitações e deficiências da TCT, a Teoria de Resposta ao Item (TRI). Este capítulo, com as devidas adaptações, comporá o material oferecido aos colegas professores como um guia para compreender essas avaliações de larga escala, entre as quais pode-se enquadrar o ENEM. Este material para o professor encontra-se no Anexo 1 deste trabalho.

Enfim, com base nas discussões apresentadas nos Capítulos 2 e 3, os Capítulos 4 e 5 trazem o coração do trabalho: os resultados da pesquisa desenvolvida. No Capítulo 4, utilizando-se das respostas de alunos em um questionário de avaliação de conceitos de física térmica, constrói-se para cada uma das questões (ou itens) uma curva característica do item (ICC). A partir destas, que se constituem num instrumento de análise utilizado na TRI, procede-se à análise dos resultados obtidos e faz-se uma discussão sobre a propriedade da utilização das mesmas em um instrumento de avaliação qualquer.

No Capítulo 5, faz-se um estudo, qualitativo e semi-quantitativo, sobre as características constituintes das provas do Exame Nacional do Ensino Médio a partir de 2009, ano em que a proposta do Exame é modificada e a

metodologia utilizada para análise dos resultados das respostas dos examinandos passa a ser feita por TRI. O desdobramento natural deste trabalho é analisar os microdados, ainda não disponíveis, dos exames do ENEM de 2009 a 2011 a partir da análise qualitativa proposta.

Finalizando, no Capítulo 6 apresentam-se as considerações finais, certo de que apossar-se e dominar os conceitos envolvidos com as avaliações de larga escala pode fazer a diferença no trabalho dos professores, especialmente do ensino médio, pois essas avaliações revelam muito a respeito da educação fornecida pelas escolas, educação entendida como o resultado final do processo de ensino. Professores capazes de interpretar os resultados provavelmente serão capazes de melhorar suas estratégias didáticas, seus materiais e sua visão de mundo. O trabalho se propôs a ser uma etapa na construção deste conhecimento dos professores.

CAPÍTULO 2 – AVALIAÇÕES DE LARGA ESCALA

Neste capítulo apresenta-se uma breve discussão sobre os processos de avaliação em educação. Inicia-se com a apresentação das principais funções dessas avaliações: o aspecto formativo, o somativo e o somativo associado a processos de responsabilização (“accountability”). Em seguida, faz-se um levantamento dos processos de avaliação de larga escala dos sistemas educacionais de caráter internacional e os existentes no Brasil. O foco da discussão é a avaliação em ciências.

2.1- Aspectos Gerais sobre Avaliação de Aprendizagem

Os processos de avaliação, com todas as suas particularidades, são fontes geradoras de sentimentos que passam do amor ao ódio, e ocupam o centro das atenções em qualquer processo de ensino e aprendizagem.

No planejamento e construção dos projetos político-pedagógicos, na seleção dos conteúdos a serem abordados, no planejamento e reestruturação das aulas, na classificação, hierarquização e seleção dos estudantes, a avaliação desempenha um papel fundamental, possibilitando a obtenção de informações para análise, a análise propriamente dita dos dados coletados e o direcionamento das ações a serem tomadas.

Mas, o que é avaliar? Quais os tipos de avaliação? Quais os métodos e processos disponíveis? O que pensam de tal prática e como a utilizam os professores e alunos? As respostas para estas questões estão longe de alcançar um consenso, seja entre os pesquisadores da área, seja entre os que fazem uso desses processos no desenvolvimento de suas atividades profissionais [Perrenoud 1999].

Partindo do princípio que a compreensão do processo avaliativo depende fundamentalmente do contexto histórico, social e político em que esse processo ocorre [Black 1998], definir o que é avaliar ou o que é uma avaliação

está diretamente ligado ao saber o que se quer avaliar. Segundo Black [Black 1998], avaliação é

“(...) um termo genérico que engloba todos os métodos usualmente utilizados para avaliar o desempenho de um aluno ou grupo. Quando se fala de avaliar, pode-se considerar uma avaliação ampla que inclui muitas fontes de evidências e muitos aspectos do conhecimento, das habilidades e das atitudes do aluno, ou então, se o objetivo é avaliação de um aspecto do estudante em particular, a avaliação pode ser entendida como qualquer método ou procedimento, formal ou informal, para a produção de informações sobre esses alunos.” (p. 5) ¹

De forma também muito abrangente, Perrenoud [Perrenoud 1999] apresenta a seguinte definição para avaliação

“Avaliar é – cedo ou tarde – criar hierarquias de excelência, em função das quais se decidirão a progressão no curso seguido, a seleção no início do secundário, a orientação para diversos tipos de estudos, a certificação antes da entrada no mercado de trabalho e, frequentemente, a contratação. Avaliar é também privilegiar um modo de estar em aula e no mundo, valorizar formas e normas de excelência, definir um aluno modelo, aplicado, dócil para uns, imaginativo e autônomo para outros...” (p.9)

Nessas definições, a finalidade da avaliação está apresentada como o ponto fundamental para a definição do processo; citando Black [Black 1998]:

“ É importante que a seleção do processo de avaliação e sua utilização correspondam ao fim específico que se destina sendo que, desde o início desse processo, já se deve fazer a distinção entre os objetivos e os instrumentos e procedimentos que podem ser utilizados. Por exemplo, as questões de um mesmo teste podem ser utilizadas para finalidades completamente diferentes, e, em sentido contrário, uma determinada finalidade pode ser obtida através da combinação dos resultados de uma gama de tipos de avaliações diferentes. Além disso, as avaliações podem ser realizadas por diferentes agentes, desde o professor até um comitê internacional de pesquisa.” (p. 24) ²

¹ Original: “A general term embracing all methods customarily used to appraise performance of individual pupil or group. It may refer to a broad appraisal including many sources of evidence and many aspects of a pupil’s knowledge, understanding, skills and attitudes; or to a particular occasion or instrument. An assessment instrument may be any method or procedure, formal or informal, for producing information about pupils.” (p. 5)

² Original: “It is important to match the selection and use of assessment methods to the particular purpose which the assessment is meant to serve and a distinction has to be made at the outset between the purposes and the instruments and procedures that might be used. For example, the same test questions

Em uma escola, a avaliação é de suma importância para a construção e manutenção de seu projeto pedagógico e para o desenvolvimento das intenções educativas desse projeto. Em outras palavras, a avaliação está a serviço do projeto educacional, é parte integrante do mesmo e partilha de seus princípios fundamentais.

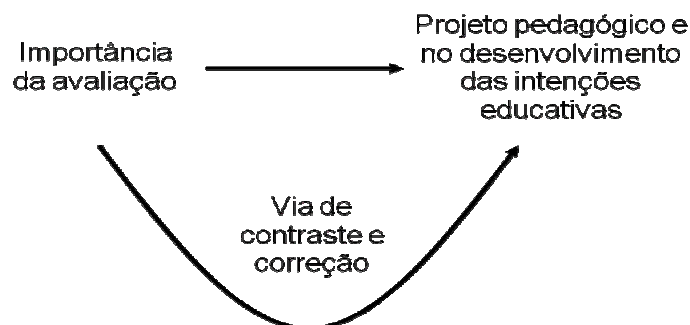


Figura 2.1. Importância da avaliação no processo pedagógico.

Pode-se fazer uma classificação dentre as diversas finalidades possíveis para um processo avaliativo em três grupos [Black 1998]: o primeiro, usualmente chamado de avaliação formativa, que contribui para o desenvolvimento e aprendizagem do aluno; o segundo, avaliação somativa, que seleciona e hierarquiza os que dele participam; e o terceiro, avaliação com responsabilização³, que satisfaz as demandas para a responsabilização pública. Alguns autores (como, por exemplo, Perrenoud) não separam a avaliação em três, e sim em dois grupos: avaliação somativa e avaliação formativa. Apresenta-se a seguir uma discussão mais detalhada das características de cada um desses tipos de avaliação.

2.1.1. Avaliação Formativa

Considera-se importante que o professor obtenha, de forma contínua, informações sobre o que seu aluno sabe e que estratégias ele deve usar com

may be used for quite different purposes, and conversely, a single purpose might be served by combining the results obtained from a range of different types of assessment. Furthermore, assessments may be carried out by many different agencies, from the teacher in the classroom to a committee mounting an international survey.” (p. 24)

³ O termo é uma tradução às vezes utilizada para a palavra inglesa “accountability”.

este estudante para que ele processe e compreenda os novos conceitos apresentados na atividade escolar.

Esse processo de avaliação, que costuma ser denominado de formativo, possui o objetivo de subsidiar o processo de ensino-aprendizagem, caracterizando-se pela obtenção constante de informações que permitem intervenções e contínuas modificações nas atividades de ensino do professor.

Nesse processo avaliativo, o retorno (“feedback”) de sala de aula é essencial para a condução do ensino. Nenhum sistema, mecânico ou social, pode se ajustar ou se adaptar, à medida que executa uma tarefa, sem receber informações de seu processo operacional [Black, 1998]. O retorno, através de atividades avaliativas, possibilita que os processos de ensino e aprendizagem variem em ritmo e estilo de acordo com as necessidades dos estudantes [Black 1998]:

“(...) os professores precisam ir devagar ou repetir o que foi trabalhado em sala quando percebem as dificuldades de seus alunos. Eles também precisam mudar a forma como ensinam quando obtêm evidências que indicam que alguns captaram as ideias e podem progredir enquanto outros se atrapalharam e não conseguem ir em frente.” (p. 25)⁴

Em outras palavras, avaliação formativa é aquela que possui a finalidade de fornecer informações acerca do desenvolvimento de um processo de ensino e aprendizagem, para que o professor possa fazer as intervenções e mudanças de acordo com as características de seus alunos. Este tipo de avaliação caracteriza-se por não possuir uma finalidade probatória, mas sim a função de inquirir, harmonizar, apoiar, orientar, reforçar, corrigir, melhorar o processo de ensino e aprendizagem, preocupando-se com a ação de formar o aluno. O fato de informar continuamente ao professor sobre o desenvolvimento do processo de ensino e aprendizagem em sala de aula e aos alunos sobre seus domínios e deficiências faz desse tipo de avaliação uma contribuição para a melhoria da aprendizagem.

⁴ Original: “(...) teachers need to go slowly, or repeat what has been done, when difficulties of their pupils become apparent. They need also to differentiate their teaching as they collect evidence that some have grasped ideas and want to go ahead, whereas others are trapped in confusions so that they are unable to go ahead.” (p. 25)

2.1.2. Avaliação Somativa

O termo avaliação somativa implica numa visualização generalizada do conhecimento que é trazido pelo estudante da etapa anterior para a que a sucede. Esta visualização pode ser adquirida através de um acúmulo de evidências coletadas ao longo do tempo ou de processos avaliativos aplicados no final de cada etapa, atribuindo ao aluno uma nota (normalmente um valor numérico) que atestará seu sucesso ou fracasso no processo de formação escolar.

A escolha do que é mais importante nesse processo avaliativo, em geral, constitui-se em um problema. Isto acontece em virtude do fato de que essa escolha, na maioria das vezes, tem caráter pessoal, dependendo do tipo de formação e afinidades do professor que a utiliza [Black, 1998].

De uma forma generalizada e abrangente, as finalidades do processo de avaliação somativa são [Perrenoud, 1999]:

- a) a criação de níveis de excelência;
- b) o surgimento de uma negociação entre o professor e o aluno, fazendo com que a avaliação torne-se um instrumento de autoridade e controle de disciplina;
- c) a informação para os alunos e seus pais sobre a possibilidade de fracasso ou sucesso nos graus de aprendizagem estabelecidos pelos sistemas educacionais;
- d) um modo de regulação da divisão vertical do trabalho pedagógico, isto é, a avaliação serve para controlar o trabalho dos alunos e, simultaneamente, para gerir os fluxos.

2.1.3. Avaliação somativa com responsabilização (“accountability”)

A busca por um lugar competitivo no mercado, associado aos grandes investimentos necessários aos sistemas educacionais, têm levado os governos a se preocupar cada vez mais com os resultados apresentados por estes sistemas [Brooke, 2006]. Neste contexto, surge uma nova função para avaliações somativas, uma função que extrapola os “muros” da escola, que

permite comparações entre as instituições escolares e que fornece diagnósticos para os sistemas; em suma, apresenta um caráter de prestação de contas à sociedade e aos financiadores da escola.

Este sistema de responsabilização a partir de avaliações somativas, de uma forma geral, é constituída de quatro ingredientes básicos [Brooke, 2006]:

- “1) a decisão por parte das autoridades de tornar públicas as diferenças de nível de desempenho das escolas (ingrediente autoridade);*
- 2) o uso de testes ou procedimentos padronizados para fornecer este tipo de informação (ingrediente informação);*
- 3) os critérios para analisar esta informação e para determinar quais escolas têm melhor desempenho (ingrediente padrões);*
- 4) os critérios para aplicação de incentivos ou sanções conforme os padrões estabelecido (ingrediente consequências).” (p. 380)*

Com isso, tornam públicas as informações sobre os trabalhos das escolas e são atribuídos aos diretores e outros membros da escola uma corresponsabilidade pelo nível de desempenho alcançado pela instituição [Brooke, 2006]. No Brasil, pode-se citar como exemplo deste tipo de avaliação o *“Prêmio Educacional Escola do Novo Milênio – Educação Básica de Qualidade no Ceará”*, o *“Programa Nova Escola da Secretaria de Educação do Estado do Rio de Janeiro”*, dentre vários outros.

2.2- Aspectos das Avaliações de Larga Escala

Avaliações de larga escala constituem um tipo de avaliação educacional que tem como objetivo a obtenção de dados e a realização de análises que possibilitem diagnosticar a situação da educação, fornecendo subsídios para a implementação, manutenção e reformulação de políticas educacionais. Para isso, elas são realizadas em “larga escala”, isto é, com um número muito grande de testes aplicados ou com a utilização de algum tipo de amostragem estatística. A possibilidade de monitoramento contínuo oferecido por este tipo de avaliação permite detectar os benefícios e malefícios que decorrem de políticas adotadas [Klein e Fontanive 1995]. Segundo Britton e Schneider [Britton e Schneider 2007]

“Historicamente, as avaliações de larga escala são utilizadas para fazer comparações internacionais, políticas de responsabilização (“accountability”), reportar resultados para fora da sala de aula, promover a qualificação para entrada nas instituições de ensino e para avaliar programas. Portanto essa denominação engloba uma grande variedade de finalidades, formatos e outras características. Tem-se, então, que ao se discutir uma avaliação de larga escala é importante se considerar as especificidades dessas avaliações para não se cometer generalizações inapropriadas.” (p. 1007)⁵

Uma das características das avaliações de larga escala é que elas em geral não se destinam a analisar e fornecer dados de alunos ou escolas individualmente, e sim analisar populações e subpopulações de estudantes em diferentes níveis, muitas vezes permitindo o acompanhamento destas ao longo do tempo.

Nos últimos anos, observa-se que o número de avaliações de larga escala tem crescido substancialmente em todos os níveis do sistema educacional. Estas avaliações, em geral, abordam a aprendizagem de linguagens e matemática. No entanto, desde 2000, muitas comparações vêm sendo feitas, na área de ciências, por organizações internacionais que patrocinam comparações da educação básica entre vários países. Observa-se que a adesão dos países a essas avaliações vem crescendo [Britton e Schneider 2007].

No Brasil, além de funcionar como um instrumento de diagnóstico e obtenção de dados, uma das avaliações aplicada nacionalmente, o Exame Nacional do Ensino Médio (ENEM), passou, também, a fornecer a certificação de conclusão do ensino médio, a possibilidade de ingresso em um grande número de universidades públicas e privadas e a possibilidade de acesso ao financiamento estudantil.

A compreensão e entendimento desse tipo de avaliação tornam-se fundamentais, principalmente para os professores, a partir do momento que, ao servir de base para as ações das políticas públicas em educação, passam a determinar, direta e indiretamente, o currículo a ser ensinado nas escolas, as

⁵ Original: “Historically, these assessments are for purposes such as international comparisons, accountability, summative assessment for reporting to others outside the classroom, qualification for entry to college, and evaluation of programs. Thus, the term “large-scale assessments” encompasses quite a variety of assessment that differ in their purposes, formats, and other features. It’s important to consider the specifics of those features to avoid inappropriate generalizations.” (p. 1007)

cargas horárias das disciplinas e, finalmente, o perfil dos alunos que ingressam nas universidades.

A seguir, apresenta-se um levantamento das avaliações de larga escala de caráter internacional sobre sistemas nacionais, seguida pelas avaliações nacionais em nosso país.

2.3- Avaliações Internacionais de Larga Escala

Um levantamento das avaliações de larga escala é apresentado por Britton e Schneider [Britton e Schneider 2007]. A Tabela 2.1 corresponde à tabela apresentada por esses autores, atualizada até o ano de 2011.

Tabela 2.1. As principais avaliações internacionais de larga escala.

Ano	Organização Responsável	Sigla	Conteúdos Abordados	Países Participantes
1969-70	IEA	FISS	Ciências	19
1983-86	IEA	SISS	Ciências	17
1988	ETS	IEAP	Ciências e Matemática	6
1992	ETS	IEAP	Ciências e Matemática	20
1994-95	IEA	TIMSS	Ciências e Matemática	42
1999	IEA	TIMSS-R	Ciências, Matemática	39
2000	OECD	PISA	Leitura (ciências/matemática)	39
2003	IEA	TIMSS	Ciências e Matemática	48
2003	OECD	PISA	Matemática (leitura/ciências)	41
2006	OECD	PISA	Ciências (leitura/Matemática)	57
2007	IEA	TIMSS	Ciências e Matemática	50
2009	OECD	PISA	Leitura (Matemática/Ciências)	65
2011	IEA	TIMSS	Ciências e Matemática	77

Legenda: IEA = Associação Internacional para a Avaliação do Desenvolvimento Educacional (International Association for the Evaluation of Educational Achievement), ETS = Serviço de Testes Educacionais (Educational Testing Service), OECD = Organização para Cooperação Econômica e Desenvolvimento (Organization for Economic Co-Operation and Development), FISS= Primeiro Estudo Internacional em Ciências (First International Science Study), SISS = Segundo Estudo Internacional em Ciências (Second International Science Study), TIMSS= Tendências nos Estudos de Matemática e Ciências (Trends in Mathematics and Science Study), PISA = Programa para Avaliação Internacional dos Estudantes (Programme for International Student Assessment).

Observa-se desta tabela que as avaliações de larga escala em ciências e matemática aplicadas por organizações internacionais estão ficando maiores e mais frequentes. Em seguida, apresenta-se com alguns detalhes algumas dessas avaliações.

2.3.1. PIRLS, TIMMS – A IEA

Fundada em 1958, em Amsterdam, a IEA (*The International Association for the Evaluation of Student Achievement*) é uma associação privada, sem fins lucrativos, que realiza estudos comparativos, a nível internacional, da qualidade de ensino das escolas. A IEA inclui em seus membros cerca de 66 países que participam de sua organização e tem desenvolvido trabalhos com mais de 90 sistemas educacionais em vários outros países. Maiores informações sobre as avaliações desenvolvidas podem ser encontradas na página <http://www.iea.nl/>.

O foco do trabalho realizado pela IEA está na condução de um estudo sobre o desempenho dos estudantes em áreas específicas como matemática, ciências e leitura. Através desses projetos de pesquisa comparativos e de avaliação, a IEA tem como objetivos [Wagemaker, 2008]:

- Fornecer referências internacionais que possam ajudar na construção de políticas educacionais identificando os pontos fortes e fracos dos sistemas educacionais em estudo.
- Fornecer informações de alta qualidade para que os elaboradores e gestores de políticas educacionais compreendam os fatores escolares e não escolares que influenciam o ensino e aprendizagem.
- Fornecer dados de alta qualidade que sirvam como fonte para a identificação de áreas de preocupação e ação e que possibilitem a preparação e avaliação de reformas educacionais.
- Desenvolver e melhorar a capacidade dos sistemas educacionais em engajar-se em estratégias nacionais para monitoramento e melhoria.

- Contribuir para o desenvolvimento da comunidade mundial de pesquisadores em avaliação educacional.

Como exemplo de estudos promovidos pela IEA tem-se o “Progress in International Reading Study” (PIRLS) e o “Trends in Mathematics and Science Study” (TIMSS), que avaliam os resultados de aprendizagem dos alunos que atingiram determinadas séries, independente da idade dos mesmos nestas etapas. O argumento para essa escolha baseia-se na crença que o aluno aprende através do ensino e não por tornar-se mais velho. Estes estudos estão fundamentalmente preocupados com a oportunidade de aprender, isto é, com a compreensão das ligações entre o que é descrito como o currículo pretendido (o que é ditado pela política educacional), como o currículo implementado (que é ensinado nas escolas) e como o currículo alcançado (o que os alunos aprendem).

Entre os anos de 1959 e 1962, um estudo-piloto envolvendo doze países avaliou cinco áreas: ciências, matemática, interpretação de textos, geografia e habilidades não-verbais. Os estudos da IEA são conhecidos por avaliar o conteúdo científico, mas são também planejados para avaliar a expectativa de desempenho dos estudantes em ciências e as perspectivas para a ciência, como pôde ser observado no TIMSS, em 1995 [Britton e Schneider 2007].

O FISS (First International Science Study) foi conduzido pela IEA entre 1966 e 1973. Participaram do estudo três grupos de estudantes, separados por suas idades: dez anos, catorze anos e estudantes do último ano do ensino secundário. Neste estudo, apenas o grupo dos estudantes de dez anos de idade foram avaliados em ciências da terra, enquanto os dois grupos de estudantes mais velhos foram avaliados em ciências da natureza e metodologia científica. Todos os três grupos foram avaliados nas áreas de biologia, química e física. Dezenove sistemas educacionais nacionais participaram nas avaliações dos dois grupos de estudantes mais velhos e dezessete participaram nas avaliações do grupo de estudantes de dez anos de idade.

O SISS (Second International Science Study) foi conduzido entre 1983 e 1986 e procurou aplicar o que foi aprendido com o FISS e com o Second International Mathematics Study (SIMS), que havia acabado de ser realizado. Havia interesse em examinar as diferenças nas avaliações entre os países e reunir informações que pudessem sugerir explicações para estas diferenças. O número de países que participaram das avaliações novamente variou conforme a idade dos grupos de estudantes. Quinze sistemas educacionais nacionais participaram nas avaliações dos estudantes de dez anos de idade, dezessete nas dos estudantes de catorze anos e outros catorze sistemas nas avaliações dos estudantes do último ano do ensino secundário. Para as propostas da avaliação, os estudantes mais velhos foram divididos em quatro grupos: aqueles que estavam estudando biologia, os que estudavam química, os que estudavam física e aqueles que não estavam estudando nenhuma ciência na época da avaliação. Keeves [Keeves 1992] discutiu em um artigo as transformações nas avaliações de dez países que participaram tanto do FISS quanto do SISS, durante os catorze anos que separaram os dois estudos. Em seguida, a IEA organizou o PIRLS e o TIMSS, discutidos em mais detalhe.

Os estudos da IEA são uma fonte de dados importante para aqueles que trabalham para melhorar a aprendizagem dos alunos a nível internacional, nacional e local. Ao tratar de uma vasta gama de tópicos e assuntos, os estudos contribuem para uma compreensão profunda dos processos educacionais em cada país num amplo contexto internacional. Além disso, o ciclo de estudos fornece aos países a oportunidade de medir o progresso na na compreensão em matemática, em ciências e em leitura. O ciclo de estudos também permite o monitoramento de mudanças na implementação da política educacional e na identificação de novos temas relevantes para os esforços de reforma.

2.3.2. PIRLS - Progresso no Letramento de Leitura

O PIRLS (Progress in Reading Literacy) é concebido como um estudo de tendência, em um ciclo de cinco anos, destinado a estudantes da 4ª série que, em muitos países, corresponde ao momento em que esses estão fazendo

a transição do aprender a ler para o momento em que se usa a leitura para aprender. Para que estejam em condições de fazer a prova, a média de idade da série aplicada não deve ser menor que 9,5 anos.

O PIRLS define como “leitura” [Wagemaker 2008]:

“A habilidade de entender e usar a linguagem escrita exigida pela sociedade ou valorizados pelo indivíduo. Jovens leitores podem construir significados de uma variedade de textos. Eles lêem para aprender, participar de comunidades de leitores e para se divertir (PIRLS Framework 2004).” (p. 5)⁶

Visando proporcionar aos educadores, professores e formuladores de políticas uma visão mais abrangente das competências e habilidades dos estudantes, o exame avalia o conteúdo e os domínios cognitivos relativos às finalidades da leitura e aos processos de compreensão de textos.

Além dessas medidas, centrais para se entender o que os alunos sabem e podem fazer, o PIRLS obtém quantidades significativas de informação através de aplicação de questionários sobre o contexto nacional e sobre o currículo, aplicados aos diretores das escolas, aos professores, aos estudantes e seus pais. Estes questionários fornecem importantes dados para que se saiba como a escola, a casa e os recursos disponíveis favorecem o hábito da leitura, além de permitir a compreensão das ligações existentes entre a política a nível nacional e a sala de aula. Com esses dados, é também possível a comparação entre os países que participam do processo de avaliação.

2.3.3. TIMSS - Tendências Internacionais nos Estudos de Matemática e Ciências

Assim como o PIRLS, o TIMSS (Trends in International Mathematics and Science Study) é um estudo de tendência elaborado para ciclos de quatro anos. Ele é aplicado na 4ª e 8ª séries do ensino fundamental (5º e 9º anos no Brasil). O ciclo de quatro anos, além de permitir o estudo dos grupos ao longo

⁶ Original: “The ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety texts. They read to learn, to participate in communities of readers, and for enjoyment.” (PIRLS Framework 2004, p. 5)

do tempo, possibilita também a análise de mudanças de desempenho desse grupo quatro anos depois. Aplicado em 1995, 1999, 2003 e 2007, teve o seu quinto ciclo completado em 2011.

O TIMSS, além de fornecer uma medida da aprendizagem em matemática e ciências em dois níveis consecutivos de ensino, possibilita também uma visão geral dos conteúdos e domínios cognitivos que orientam o desenvolvimento da avaliação e servem de base para relatórios mais detalhados.

O TIMSS 2007 fornece um modelo para a avaliação da 4ª e 8ª séries em matemática e ciências, cujos objetivos nos domínios de conteúdo estão explicitados nas Tabelas 2.2 e 2.3.

Tabela 2.2. Distribuição percentual dos objetivos em domínio de conteúdo da avaliação de matemática do TIMSS 2007 na 4ª série.

Domínios de Conteúdo – 4ª Série	
Números	50%
Formas geométricas e medidas	35%
Apresentação de informação	15%

Tabela 2.3. Distribuição percentual dos objetivos em domínio de conteúdo da avaliação de matemática do TIMSS 2007 na 8ª série.

Domínios de Conteúdo – 8ª Série	
Números	30%
Álgebra	30%
Geometria	20%
Informação e Probabilidade	20%

Além do domínio de conteúdo, são exigidos dos estudantes uma gama de domínios cognitivos para que possam responder corretamente as questões propostas no TIMSS. A distribuição percentual desses domínios encontram-se especificadas na Tabela 2.4.

Tabela 2.4. Distribuição percentual dos objetivos dos domínios cognitivos da avaliação de matemática do TIMSS 2007 nas 4ª e 8ª séries.

Domínios Cognitivos		
	4ª Série	8ª Série
Conhecimento	40%	35%
Aplicação	40%	40%
Raciocínio	20%	25%

Construindo-se as mesmas distribuições de domínios de conteúdo e domínios cognitivos para a 4ª e 8ª séries na área de ciências, tem-se:

Tabela 2.5. Distribuição percentual dos objetivos em domínio de conteúdo da avaliação de ciências do TIMSS 2007 na 4ª série

Domínios de Conteúdo – 4ª Série	
Ciência da Vida	45%
Ciências Físicas	35%
Ciências da Terra	20%

Tabela 2.6. Distribuição percentual dos objetivos em domínio de conteúdo da avaliação de ciências do TIMSS 2007 na 8ª série.

Domínios de Conteúdo – 8ª Série	
Biologia	35%
Química	20%
Física	25%
Ciências da Terra	20%

Tabela 2.7. Distribuição percentual dos objetivos dos domínios cognitivos da avaliação de ciências do TIMSS 2007 nas 4ª e 8ª séries.

Domínios Cognitivos		
	4ª Série	8ª Série
Conhecimento	40%	30%
Aplicação	35%	35%
Raciocínio	25%	35%

Assim como no PIRLS, a aprendizagem de matemática e ciências também ocorre dentro de um contexto, e vários fatores, como o tipo de escola, os recursos disponíveis, as práticas educacionais, as atitudes dos alunos, a participação dos pais e o ambiente familiar, interferem na aprendizagem do aluno. Para avaliar a influência desses fatores, são aplicados questionários envolvendo perguntas sobre o contexto nacional, sobre os currículos adotados, aos diretores das escolas, aos professores e aos estudantes envolvidos.

Em 2011, 77 países participaram do TIMSS; mais de 20.000 estudantes em mais de 1000 escolas de todo os Estados Unidos fizeram a avaliação, juntando-se a quase 500 mil alunos em todo o mundo.

2.3.4. PISA – Programa Internacional de Avaliação dos Estudantes

O PISA (Programme for International Student Assessment) é o programa Internacional de Avaliação de Estudantes, organizado pela OCDE (Organização para Cooperação e Desenvolvimento Econômico). Constituído em ciclos de três anos, já foi aplicado nos anos de 2000, 2003, 2006 e 2009 em diversos países da Comunidade Europeia e em países convidados, como o Brasil.

Nesse programa, são avaliadas as capacidade de analisar, raciocinar e refletir sobre conhecimentos e experiências na áreas de Leitura, Matemática e Ciências, sendo que em cada uma de suas edições, uma dessas três áreas é focada, como mostrado na Tabela 2.8.

Tabela 2.8. Área de foco do Pisa nos anos de sua aplicação.

Ano	Área de Foco
2000	Leitura
2003	Matemática
2006	Ciências
2009	Leitura

A aplicação deste teste é feita por amostragem, atingindo em geral entre 4500 a 10000 estudantes em cada um dos países participantes. A seleção dos participantes, no Brasil, é feita com base no Censo Escolar. Os alunos participantes são os que têm 15 anos e estão pelo menos cursando a 6ª série. O programa avalia a aquisição de alguns dos conhecimentos e habilidades considerados essenciais para a plena participação na sociedade atual:

“O objetivo principal do PISA é produzir indicadores que contribuam para a discussão da qualidade da educação ministrada nos países participantes, de modo a subsidiar políticas de melhoria da educação básica. A avaliação procura verificar até que ponto as escolas de cada país participante estão preparando seus jovens para exercerem o papel de cidadãos na sociedade contemporânea.” (Portal INEP, 2011).⁷

O PISA utiliza o conceito de letramento para definir a amplitude de conhecimentos, as habilidades e as competências a serem avaliados no exame. Para cada uma das três áreas, o conceito é definido segundo [Bonamino, Coscarelli e Franco 2002; Barroso e Franco 2008]

“Letramento em Leitura: compreender, usar e refletir a respeito de textos escritos, para atingir o objetivo individual, para desenvolver o conhecimento e potencial individual, e para participar da sociedade.”⁸

“Letramento em Matemática: capacidade de identificar e compreender o papel que a matemática desempenha no mundo, de fazer julgamentos bem fundamentados e usar e se relacionar com a matemática de maneira que atinjam as necessidades da vida individual de um cidadão construtivo, preocupado e reflexivo.”⁹

“Letramento em Ciências: capacidade de usar o conhecimento científico para identificar questões e chegar a conclusões baseadas em evidências para entender e ajudar a tomar decisões a respeito do mundo e as mudanças causadas a ele pela atividade humana.”¹⁰

⁷ Portal INEP, 2011, disponível em www.inep.gov.br, consultado em 29/06/2011.

⁸ Original: “Reading literacy is understanding, using, and reflecting on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society.”

⁹ Original: “Mathematical literacy is an individual’s capacity to identify and understand the role that mathematics plays in the world, to make well founded judgments and to use and engage with mathematics in ways that meet the needs of that individual’s life as a constructive, concerned and reflective citizen.”

¹⁰ Original: “Scientific literacy is the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.”

A última destas três definições, o “letramento científico” compreende três aspectos que devem ser avaliados de forma combinada: os conteúdos ou conhecimentos científicos, os processos científicos e as situações ou contextos nos quais os conhecimentos ou processos são avaliados.

Esta avaliação é feita, na área de conteúdos, de uma seleção dos campos da ciência física, química, biológica, da Terra e do espaço, de acordo com os seguintes critérios [Barroso e Franco 2008]:

- Importância para situações cotidianas;
- O conhecimento e as áreas de aplicação sejam de relevância duradoura para (pelo menos) a próxima década;
- O conhecimento exigido possa ser combinado com os processos científicos selecionados.

Ainda dentro destes critérios, intensamente relacionado aos conteúdos, na avaliação dos aspectos relacionados aos processos, tem-se:

- Observação das habilidades e compreensão necessárias para coletar e interpretar evidências do mundo ao nosso redor;
- Obtenção de conclusões, através do uso de conhecimentos científicos, após a coleta e interpretação de dados.

Finalmente, na avaliação do aspecto de situações ou contexto, tem-se

- Verificação da capacidade de aplicação do conhecimento científico selecionado;
- Uso de processos científicos em situações importantes refletindo o mundo real e envolvendo as ideias da ciência.

Além das avaliações das competências nas áreas específicas de Leitura, Matemática e Ciências, o programa, através da aplicação de questionários, coleta dados para elaboração de indicadores que permitam relacionar o desempenho apresentado pelos estudantes ao contexto que pertencem, delineando-se as variáveis demográficas, sociais, econômicas e educacionais específicas de cada grupo.

A característica do PISA que o distingue do TIMSS é a natureza da ciência que cada um visa avaliar, interferindo assim na natureza dos itens de teste em cada um deles. Enquanto TIMSS se concentra principalmente no conteúdo formal das disciplinas científicas, o PISA enfatiza a aplicação da ciência em contextos de vida real [Britton e Schneider, 2007].

2.4- Avaliações de Larga Escala Nacionais

Como se pode observar, as avaliações de larga escala já vêm sendo utilizadas ao longo do tempo em processos internacionais de avaliação. Cabe agora, então, uma discussão sobre as avaliações de larga escala no Brasil.

2.4.1. O SAEB e a Prova Brasil

Tendo sua primeira aplicação no ano de 1990, o Sistema de Avaliação da Educação Básica (SAEB) foi desenvolvido com o objetivo de fornecer um diagnóstico do sistema educacional brasileiro e de verificar a interferência de alguns fatores como as condições socioeconômicas dos alunos, e as condições físicas e recursos das escolas no desempenho desses alunos, fornecendo, assim, um indicativo da qualidade de ensino ofertado no sistema educacional brasileiro.

Em sua primeira edição, o SAEB contou com a participação de uma amostra de escolas públicas da rede urbana que ofertavam as 1^a, 3^a, 5^a e 7^a séries do Ensino Fundamental. Foram feitas avaliações em língua portuguesa, matemática e ciências, sendo que as 5^a e 7^a séries também foram avaliadas em redação. Esse formato foi mantido até 1993.

Em 1995, ocorre a adoção da Teoria de Resposta ao Item para a construção do teste e análise dos resultados dessa avaliação. Essa inovação permitiu que a partir deste momento fossem comparados os resultados com as posteriores edições do SAEB. Nesse ano foram avaliadas as séries finais do fundamental (5^a e 9^o anos) e do ensino médio (3^a série), sendo que só foram

aplicados testes de língua portuguesa e matemática. Além das escolas da rede pública, também nessa edição, foi acrescentada uma amostra da rede privada.

Aplicado de dois em dois anos, a partir de 1995, as séries e áreas curriculares avaliadas podem ser encontradas na Tabela 2.9.

Tabela 2.9. Cronologia de aplicação Saeb.

Ano	Séries Avaliadas	Amostra	Área Curricular de Avaliação
1995	4ª e 9ª séries do E.F. e 3ª série do E.M.	rede pública e privada	língua portuguesa e matemática
1997 e 1999	4ª e 9ª séries do E.F.	rede pública e privada	língua portuguesa, matemática e ciências
	3ª série do E.M.		língua portuguesa, matemática, ciências, história e geografia
2001, 2003, 2005, 2007, e 2009	4ª e 9ª séries do E.F.	rede pública e privada	língua portuguesa e matemática
	3ª série do E.M.		

Também são aplicados questionários socioeconômicos aos alunos participantes e à comunidade escolar.

As edições de 1990 e 2003 tiveram caráter amostral, sendo aplicadas em grupos de escolas sorteadas, o que possibilitou gerar resultados em nível de Brasil, Região e Estados.

Em 2005, a Portaria Ministerial nº 931 de 21 de março de 2005 promove uma reestruturação no SAEB, que passa a ser composto por duas avaliações complementares: A Avaliação Nacional da Educação Básica (ANEB) e Avaliação Nacional do Rendimento Escolar (ANRESC), conhecida como Prova Brasil.

Esta última, a Prova Brasil, foi idealizada com a finalidade de suprir as necessidades dos gestores públicos, educadores, pesquisadores e da sociedade em geral em relação ao ensino oferecido em cada município e escola. O objetivo de se obter esses dados é fornecer auxílio aos governos na distribuição e direcionamento de recursos técnicos e financeiros, no

estabelecimento de metas e implantação de políticas pedagógicas, tudo na busca da melhoria da qualidade de ensino.

A Tabela 2.10 (do Portal do INEP) apresenta uma breve comparação entre esses dois processos complementares de avaliação.

Tabela 2.10. Comparação entre SAEB e Prova Brasil

ANEB (SAEB)	ANRESC (Prova Brasil)
A primeira aplicação ocorreu em 1990.	A prova foi criada em 2005.
Alunos fazem prova de Língua Portuguesa (foco em leitura) e Matemática (foco na resolução de problemas).	A Prova Brasil avalia as habilidades em Língua Portuguesa (foco em leitura) e Matemática (foco na resolução de problemas).
Caráter amostral.	Aplicação censitária.
Estudantes da rede pública e privada.	Estudantes da rede pública municipal, estadual e federal.
Área rural e urbana.	Área rural (a partir de 2007) e urbana.
5º e 9º anos E.F e 3ª série E.M.	5º e 9º anos E.F.
A avaliação é amostral, ou seja, apenas parte dos estudantes brasileiros das séries avaliadas participam da prova. Atende aos critérios estatísticos de no mínimo 10 estudantes por turma.	A avaliação é quase universal: todos os estudantes das séries avaliadas, de todas as escolas públicas urbanas do Brasil com o mínimo de 20 alunos matriculados na série avaliada, devem fazer a prova.
Por ser amostral, oferece resultados de desempenho apenas para o Brasil, regiões e unidades da Federação.	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. Esses dados são utilizados para o cálculo do Ideb.
Foco na gestão da educação básica.	Foco na implantação de ações pedagógicas e administrativas, visando à melhoria da qualidade do ensino.

2.4.2. O Índice de Desenvolvimento da Educação Básica (IDEB)

Criado em 2007 pelo Instituto Nacional de Ensino e Pesquisa (INEP), o Índice de Desenvolvimento da Educação Básica (IDEB) reúne em um único indicador dois importantes conceitos para a qualidade da educação: o de fluxo escolar e o de médias de desempenho em avaliações.

Os dados sobre a aprovação escolar colhidos no Censo Escolar e as médias obtidas nas avaliações aplicadas pelo INEP (SAEB para as unidades da federação e para o país e a Prova Brasil para os municípios) são as informações utilizadas para o cálculo desse índice.

Tendo seus valores possíveis entre 0 e 10 e podendo ser comparado nacionalmente, o IDEB permite um equilíbrio entre as duas variáveis: fluxo escolar e aprendizagem. Qualquer distorção do sistema, como, por exemplo, a retenção de alunos para melhorar o desempenho nas provas ou a aceleração do fluxo para a melhoria do índice, provoca um desequilíbrio nesse par de variáveis que pode ser detectado e corrigido.

Além disso, funciona como ferramenta para orientação de políticas públicas em prol da qualidade da educação. Por exemplo, o Plano de Desenvolvimento da Educação (PDE) estabelece como meta para 2022 que o Brasil obtenha um Ideb de média 6,0, o que corresponderia a um sistema educacional de qualidade comparável à média dos países desenvolvidos. De acordo com o INEP,

“O Ideb é mais que um indicador estatístico. Ele nasceu como condutor de política pública pela melhoria da qualidade da educação, tanto no âmbito nacional, como nos estados, municípios e escolas. Sua composição possibilita não apenas o diagnóstico atualizado da situação educacional em todas essas esferas, mas também a projeção de metas individuais intermediárias rumo ao incremento da qualidade do ensino. As metas são exatamente isso: o caminho traçado de evolução individual dos índices, para que o Brasil atinja o patamar educacional que têm hoje a média dos países da OCDE. Em termos numéricos, isso significa evoluir da média nacional 3,8, registrada em 2005, para um Ideb igual a 6,0, na primeira fase do ensino fundamental.”¹¹

Apresentadas bienalmente de 2007 a 2011, é importante ressaltar que essas metas são diferenciadas para cada rede de escolas, que deverão melhorar seus índices contribuindo, assim, em conjunto, para que o Brasil atinja a meta almejada em 2022.

¹¹ Portal INEP, 2011, disponível em www.inep.gov.br, consultado em 29/06/2011

O cálculo do Ideb é feito a partir de duas componentes: a taxa de rendimento escolar (aprovação) e as médias de desempenho nos exames padronizados aplicados pelo INEP.

A forma geral do Ideb é dada por [Portal INEP 2011]:

$$IDEB = \frac{1}{T} \times Nota,^{12}$$

onde “T” é o número de anos que os alunos de uma escola, rede de ensino, Município, Estado ou País, em média, levam para completar uma série e a “Nota” é a média obtida na Prova Brasil para as mesmas instâncias, transformada de forma a ser expressa por um valor entre 0 e 10.

Observe que quando se tem um fluxo escolar sem problemas (dentro do previsto), o termo (1/T) assume um valor unitário e a nota do Ideb equivale à nota obtida na Prova Brasil. Em casos de maior reprovação, abandono, o mesmo termo assume valores menores que a unidade, diminuindo assim o valor do IDEB. Da mesma forma, para um dado valor do termo (1/T), quanto maior a média obtida na Prova Brasil, melhor, também, o IDEB.

2.4.3. O Exame Nacional do Ensino Médio (ENEM)

O Exame Nacional do Ensino Médio – ENEM foi criado em 1998, com o objetivo de “avaliar o desempenho do indivíduo ao término da escolaridade básica, para aferir o desenvolvimento de competências fundamentais ao exercício pleno da cidadania” (ENEM – Documento Básico 1998¹³).

O exame é realizado pelo INEP, com aplicação anual, e participação voluntária dos estudantes egressos do ensino médio. De acordo com o documento básico,

“Esse exame difere de outras avaliações já propostas pelo Ministério da Educação. Centra-se na avaliação de desempenho por competências e vincula-se a um conceito mais abrangente e estrutural da inteligência humana. O exame é constituído de uma prova única, e abrange as

¹² Portal INEP, 2011, disponível em www.inep.gov.br, consultado em 29/06/2011

¹³ INEP, ENEM - Documento Básico, 1998. Apresentado no Anexo 1. Esteve disponível na página do INEP (www.inep.gov.br) até 2008.

várias áreas de conhecimento em que se organizam as atividades pedagógicas da escolaridade básica no Brasil.” (p. 1)

No período de 1998 a 2008, o exame era composto de uma prova única contendo 63 questões objetivas de múltipla escolha e uma redação.

O desempenho dos participantes era obtido através de duas notas: a nota da prova objetiva, cujo escore correspondia ao percentual de questões respondidas de forma correta (num total de 0 a 100), e a nota da redação, também valendo 100 pontos.

O exame era elaborado com base em uma Matriz de Referência com uma concepção de conhecimento que [ENEM, Documento Básico 1998]

“pressupõe a colaboração, a complementaridade e integração entre os conteúdos das diversas áreas do conhecimento presentes nas propostas curriculares das escolas brasileiras de ensino fundamental e médio, e considera que conhecer é construir e reconstruir significados continuamente, mediante o estabelecimento de relações de múltipla natureza individuais e sociais”.(p. 4)

Mais: a definição de competência previsto neste exame é [ENEM, Documento Básico 1998]

“Competências são as modalidades estruturais da inteligência, ou melhor, ações e operações que utilizamos para estabelecer relações com e entre objetos, situações, fenômenos e pessoas que desejamos conhecer. As habilidades decorrem das competências adquiridas e referem-se ao plano imediato do ‘saber fazer’. Por meio das ações e operações, as habilidades aperfeiçoam-se e articulam-se, possibilitando nova reorganização das competências.” (p. 5)

A aplicação desta definição na construção das competências para o ENEM resultou em cinco delas, que são [ENEM, Documento Básico 1998]

*I. Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica;
II. 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. Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema;*

IV. Relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente;

V. Recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.” (p. 6)

Como o objetivo declarado menciona, o ENEM fornecia um indicador da qualidade do sistema, mas não uma mensuração precisa desta qualidade. O exame não era obrigatório para os estudantes, e também não era realizado de forma amostral. Portanto, seus resultados serviam para avaliar o desempenho individual dos candidatos e, eventualmente, para fornecer um indicador para as escolas com mais de 90% de seus concluintes participando do exame.

Aos poucos, algumas instituições de ensino superior, públicas e privadas, começaram a utilizar o resultado do ENEM como forma de classificação dos candidatos ao ingresso em seus cursos. O Programa Universidade para Todos – PROUNI, criado em 2005, utiliza os resultados do ENEM para oferecer financiamentos a estudantes de escolas privadas. Com essas iniciativas, o ENEM começou a se tornar um exame cada vez mais importante para os estudantes do ensino médio, a partir do momento em que o indicador de desempenho individual passava a ser um componente do mecanismo de acesso e financiamento do nível superior em instituições públicas e privadas.

Em 2009, o MEC toma a iniciativa de propor às instituições públicas de ensino superior um acordo através do qual seria criado o Sistema de Seleção Unificada – SiSU, com o objetivo de fazer uma classificação unificada nacional do ingresso nessas instituições. O mecanismo de classificação proposto seria o resultado do ENEM. Para isso ocorre simultaneamente uma reformulação metodológica deste exame, que passa a utilizar a metodologia da Teoria da Resposta ao Item (TRI) para atribuição de escores aos candidatos, e uma reformulação conceitual, com a mudança de sua Matriz de Referência.

Os objetivos do Novo ENEM modificaram-se¹⁴: seus resultados passaram a ser uma referência para a auto-avaliação do candidato, a compor os mecanismos de seleção para acesso aos cursos profissionalizantes, pós-médios e à Educação Superior e também para o acesso a programas governamentais. Fornecem, ainda, a certificação de jovens e adultos no nível de conclusão do ensino médio e passam a ser utilizados na avaliação do desempenho acadêmico dos ingressantes nas Instituições de Ensino Superior (como componente dos resultados do ENADE, no SINAES – Sistema Nacional de Avaliação do Ensino Superior).

Os objetivos almejados com essa reformulação, segundo o INEP, são a democratização das oportunidades de acesso às vagas federais de ensino superior, a possibilidade de promover a mobilidade acadêmica e a indução de uma reestruturação dos currículos do ensino médio.

O uso desse novo exame como processo seletivo de ingresso em seus cursos de graduação foi delegado a cada uma das universidades que, como instituição autônoma, pôde utilizá-lo das seguintes formas:

- como fase única, como o sistema de seleção unificada, informatizado e on-line;
- como primeira fase;
- combinado com o vestibular da instituição;
- como fase única para as vagas remanescentes do vestibular.

No caso da certificação do ensino médio, o candidato aprovado em todas as provas do Exame deve pleitear o seu certificado de conclusão na Secretaria Estadual de Educação ou no Instituto Federal de Educação, indicada no ato da sua inscrição. São estas instituições que ficam responsáveis por estabelecer os critérios de certificação, tendo como base o edital dos exames que indicam a nota mínima para certificação em cada área do conhecimento e em redação. Para que seja possível obter essa certificação, é necessária a idade mínima de 18 anos completos na data de realização da primeira prova do exame.

¹⁴ Portal INEP, 2011, disponível em www.inep.gov.br, consultado em 29/06/2011

O Exame em si foi totalmente reformulado, tanto em sua Matriz de Referência quanto no formato das provas. Passou a se constituir de 4 (quatro) provas objetivas, contendo cada uma 45 (quarenta e cinco) questões de múltipla escolha, e uma redação. Essas provas contemplam áreas de conhecimento e componentes curriculares de acordo com a Tabela 2.11.

A Matriz de Referência, apresentada no Anexo 2, é inteiramente nova: compõe-se de 5 eixos cognitivos, comuns a todas as áreas (dominar linguagens, compreender fenômenos, enfrentar situações-problema, construir argumentação, elaborar propostas), com uma Matriz para cada uma das áreas, com competências e habilidades listadas.

Tabela 2.11. Áreas de Conhecimento e suas respectivas Componentes Curriculares¹⁵

Área do Conhecimento	Componentes Curriculares
Ciências Humanas e suas Tecnologias.	História, Geografia, Filosofia e Sociologia.
Ciências da Natureza e suas Tecnologias.	Química, Física e Biologia.
Linguagens, Códigos e suas Tecnologias e Redação.	Língua Portuguesa, Literatura, Língua Estrangeira (Inglês ou Espanhol), Artes, Educação Física e Tecnologias da Informação e Comunicação.
Matemática e suas Tecnologias.	Matemática.

No primeiro dia do exame são realizadas as provas de Ciências Humanas e suas Tecnologias e de Ciências da Natureza e suas Tecnologias, com duração de 4 horas e 30 minutos. No segundo dia, as de Linguagens, Códigos e suas Tecnologias, Redação e Matemática e suas Tecnologias, com duração de 5 horas e 30 minutos.

O desempenho dos alunos é fornecido em cada uma das provas, a partir da aplicação da Teoria da Resposta ao Item, que fornece um escore entre 0 e 1000 para cada uma delas. A nota da Redação também é fornecida como uma nota entre 0 e 1000.

Com essas reformulações ao longo do tempo, o ENEM tornou-se uma avaliação muito procurada pelos estudantes que estão concluindo o ensino

¹⁵ Portal INEP, 2011, disponível em www.inep.gov.br, consultado em 29/06/2011

médio. O número de estudantes que participa do exame é muito grande, desde o início da utilização de seus resultados no programa PROUNI. Em 2011, o número de estudantes inscritos superou 5 milhões, e um grande número de instituições federais aderiu ao SiSU (inclusive a Universidade Federal do Rio de Janeiro).

Tudo isso faz com que o ENEM forneça indicadores relevantes sobre a qualidade da educação básica brasileira, mesmo com os vícios de origem: ele não foi proposto para ser uma avaliação de sistema, e sim para ser uma avaliação individual e, principalmente, o exame não é universal nem amostral para os estudantes do ensino médio.

2.5- Comentários finais sobre processos de Avaliação de Larga Escala

O levantamento apresentado até aqui não é exaustivo, mas fornece um quadro a respeito da crescente utilização de processos de avaliação de larga escala na educação básica. Apesar de um maior número dessas avaliações concentrarem-se nas áreas de Linguagem e Matemática, as avaliações na área de Ciências vêm aumentando. Isto está ocorrendo em função de necessidades do desenvolvimento econômico, fortemente alicerçado no desenvolvimento científico e tecnológico do mundo atual.

Diante desta situação, é importante que se faça algumas considerações a respeito dessas avaliações.

Primeiramente, a ampliação no número de avaliações de larga escala em ciências pode aumentar o tempo dedicado a essa área de conhecimento dentro das salas de aula e gerar uma orientação política para essa ação. Porém, essas implicações políticas inspiradas e resultantes dessa instrução em ciências nem sempre estão de acordo com os resultados obtidos pela pesquisa em ensino e aprendizagem da área. De acordo com Britton e Schneider [Britton e Schneider 2007]:

“O aumento das avaliações de larga escala em ciências oferecem grandes oportunidades e um substancial risco para a educação em ciências. A existência dessas avaliações contribui para que haja uma

maior atenção em relação à aprendizagem de ciências na escola. Isto é particularmente verdade para os anos iniciais do ensino fundamental, onde os professores devem escolher o que ensinar, mas também há implicações para os anos finais do ensino fundamental.” (p. 1009).¹⁶

Alicerçando esta afirmação, Britton e Schneider [Britton e Schneider 2007] citam o seguinte exemplo:

“Em seus primeiros exames nos Estados Unidos, a No Child Left Behind (NCLB) promoveu somente a avaliação nas áreas de matemática e leitura, não havendo nela nenhuma parte dedicada à área de ciências. Esse fato refletiu-se diretamente na seleção áreas trabalhadas pelo professor do ensino fundamental, que privilegiou as áreas de matemática e leitura em detrimento da área de ciências, sendo este fato de amplo conhecimento da comunidade de educação em ciências. Encontraram-se ocorrências semelhantes no ensino médio. Em mais de uma escola onde os alunos apresentavam deficiências em matemática, a carga horária dessa disciplina foi dobrada, deixando as aulas de ciências em segundo plano.(...) A partir do momento que a NCLB passa avaliar também conhecimentos em ciências, ela coloca essa área no centro do palco e ajuda a garantir que a mesma possua prioridade no currículo escolar.” (p. 1009)¹⁷

Outro ponto a se observar é que determinados tipos de avaliação podem favorecer mais efetivamente o processo de ensino e aprendizagem de ciências. No entanto, a massificação desses processos avaliativos pode direcionar o ensino e aprendizagem num rumo indesejável que, muitas vezes, se torna contraproducente no que se diz respeito ao letramento científico.

Estudos realizados utilizando diversas metodologias quantitativas e qualitativas têm demonstrado resultados bastante consistentes em relação às avaliações de larga escala [Britton e Schneider 2007]:

¹⁶ Original: “Increased large-scale science assessment offers both great opportunity and substantial risk for science education. Requiring science assessment helps ensure that science gets attention in school. This is particularly true for the elementary grades, where teachers must balance and choose the amount of attention to all subjects, but it also has implications for secondary levels of schooling.” (p. 1009)

¹⁷ Original: “In the United States, the initial NCLB requirement of assessment in mathematics and reading but not in science has had a dampening effect on elementary teacher’s inclusion of science instruction. This is common knowledge in the science education community, and the first author has encountered examples of analogous effects at the middle and even high school levels during field work in recent years. He encountered more than one high school where students who, because they were failing at mathematics, were required to double up their number of mathematics courses. The concern is that this was being accomplished by dropping their formerly required general science classes. (...) by the year 2007, NCLB pushes science to the center stage of public attention and helps to ensure that it gets the priority it deserves in the school curriculum.” (p.1009)

- As avaliações de larga escala servem de foco para o planejamento e estruturação das aulas. Observa-se que professores e diretores prestam atenção no que é testado e adaptam seus currículos e forma de ensinar de acordo com estes;
- Os professores modelam o que é avaliado. Eles tendem a formatar as abordagens pedagógicas para que elas se reflitam numa melhoria do desempenho nos testes de alta visibilidade. Quando as avaliações governamentais são constituídas por questões de múltipla escolha, professores tendem trabalhar arduamente com essas questões nas atividades de sala de aula;
- Os resultados dos testes mostram uma melhora, pelo menos nos primeiros anos de aplicação. Neste período, quando novas avaliações são estabelecidas, a nota dos estudantes apresentam uma elevação.

No entanto, apesar desses benefícios gerados pela avaliação, outras pesquisas mostram consequências não intencionais nesse processo [Britton e Schneider, 2007]:

- A escola direciona seu trabalho de forma a focá-lo nos testes, muitas vezes não se preocupando a qualidade do ensino. Com as recompensas e incentivos dados a bons resultados nos testes, os educadores parecem priorizar o que é testado e como é testado, ao invés de se concentrar nos processos pedagógicos envolvidos na testagem;
- O que não é testado torna-se invisível. Como um corolário, focar nos testes ao invés do modelo também significa que o que não é testado tende ter menor atenção ou ser simplesmente ignorado. Tanto o domínio mais amplo das disciplinas testadas quanto os temas que são importantes mas não são testados passam a receber pouca atenção. Estudos têm revelado que testes governamentais tendem a dar pouca atenção ao pensamento complexo e à resolução de problemas, concentrando-se em níveis cognitivos mais baixos de desempenho dos

alunos. Isso provoca fortes implicações na determinação do que é ensinado, e como. Além disso, quando os testes e a qualidade de ensino não estão bem alinhados, parece claro que o teste e não a qualidade é o foco das atenções.

- O aumento da pontuação obtida nos testes parece ser inflado. Se os professores preparam seus alunos somente para os testes e não para os objetivos desse teste ou da educação, a pontuação obtida pode representar apenas um alto valor para aquele específico teste e não uma aprendizagem genuína de generalizações para outros domínios da ciência.

Uma outra questão a ser levantada está no tipo de questão utilizado nessas avaliações, de múltipla-escolha ou itens abertos. Muitas vezes, o tipo de questão utilizada em uma determinada avaliação de larga escala não é eficaz para obter respostas adequadas do amplo espectro de conhecimentos vistos como essenciais pela comunidade de educação em ciências. Por exemplo, as questões de múltipla escolha possuem limitações no que se diz respeito a avaliação da habilidade dos estudantes em conduzir investigações científicas.

No entanto, as questões de múltipla-escolha e de respostas curtas e diretas tendem a dominar as avaliações de larga escala por várias razões. Um desses motivos é o fator “tempo disponível” para a correção desses testes que, no caso dos de múltipla escolha e respostas curtas, permitem a avaliação de vários domínios do conteúdo científico nos prazos desejáveis. Também o custo com as correções dessas questões é menos dispendioso que o que se teria na correção de questões discursivas.

Na tentativa, no entanto, de corrigir essas distorções, algumas instituições tem produzido avaliações de larga escala empregando outros tipos de questões. O NAEP, por exemplo, tem incluído nos exames de ciências testes práticos desde de 1972. Em Israel e em alguns exames regionais feitos na Inglaterra tem-se utilizado, historicamente, práticas de laboratório como parte da avaliação de larga escala nos exames para o ensino médio [Britton e

Schneider 2007]. Uma outra tentativa é a construção de questões de múltipla escolha e respostas curtas mais complexas do que de rotina. Isso pode ser obtido, por exemplo, na construção de avaliações em ciências onde os distratores das questões de múltipla escolha são baseados na pesquisa das concepções não científicas ao invés de se basear exclusivamente na análise psicométrica de alto ou baixo desempenho. Tem-se ainda, que somando-se às questões de múltipla escolha e de respostas curtas, as avaliações de larga escala podem incluir questões com respostas mais extensas, mapas conceituais e atividades que baseiem-se em tarefas requeridas pela ciência, envolvendo, aí, atividades práticas ou baseadas em simulações computacionais de fenômenos ou dados.

A abrangência da avaliação de larga escala é outra questão que merece ser vista com cuidado. Essas avaliações devem buscar alcançar todos os estudantes, dando equivalentes oportunidades de sucesso a todos eles.

Um exemplo disso é que, em se tratando de testes internacionais, essa característica faz com que seja necessário ter cuidado com o fato que a simples tradução das questões para uma língua diferente não garante uma equivalência entre esses testes (o original e sua tradução). Uma tentativa de garantir essa equivalência é executar o processo de tradução inversa, do questionário traduzido de volta para a língua de origem do mesmo e verificar sua equivalência com o original. Esse fato tem feito com que desenvolvedores de avaliações de larga escala, no desenvolvimento dessas avaliações, se preocupem com o viés cultural das mesmas.

E por fim, é importante ressaltar que os resultados das avaliações de larga escala devem fornecer informações que possam ajudar as escolas, professores e alunos, na busca da qualidade do ensino. Para isso, é necessário que a qualidade de ensino e as avaliações estejam alinhados, o que não é uma tarefa fácil.

CAPÍTULO 3 – MEDIDAS E AVALIAÇÕES DE APRENDIZAGEM

Neste capítulo, faz-se uma breve revisão a respeito de processos e técnicas de medida. Inicia-se pela ideia de medição nas ciências físicas e a dificuldade de transpor esse conceito para as ciências sociais. Em seguida, apresentam-se alguns exemplos, tanto da física quanto da educação. Finalmente, chega-se ao ponto da fundamentação desse trabalho: discutem-se as técnicas tradicionais de avaliação de aprendizagem, a de medir a aprendizagem através da aplicação de testes (“provas”), e as teorias da psicometria para obter “notas”, ou escores, a partir desses testes, fazendo-se uma reflexão sobre a Teoria Clássica dos Testes (TCT) e as ideias básicas da Teoria da Resposta ao Item (TRI).

3.1- Medidas em Ciências Exatas e em Ciências Sociais

Ao trabalhar com ideias e grandezas em Física, ou em alguma das ciências ditas “exatas”, parece natural imaginar que tudo pode ser expresso em termos de grandezas mensuráveis que se relacionam por meio de equações.

No campo das ciências naturais, e para ser mais específico, na Física, “medir é um processo que nos permite atribuir um número a uma propriedade física como resultado de comparações entre quantidades semelhantes, sendo uma delas padronizada e adotada como unidade” [Alonso e Finn 1972].

Buscando-se generalizar este conceito para outras áreas, como o das ciências sociais, tem-se que [Allen e Yen 2002]:

“Medir é atribuir de forma sistemática números a indivíduos para representar propriedades desses indivíduos. Os números são atribuídos aos indivíduos segundo procedimentos cuidadosamente prescritos e reproduzíveis. Por exemplo, um teste de personalidade gera pontuações a partir do uso das mesmas instruções, perguntas, e procedimentos de pontuação para cada examinando. Essas pontuações não podem ser comparadas de forma significativa se a cada um dos examinandos forem dadas instruções diferentes ou se diferentes procedimentos de pontuação forem utilizados para estabelecê-los. Na mensuração, os números são atribuídos de forma sistemática e podem possuir formatos variados. Por exemplo, pode-se atribuir-se às pessoas com cabelos vermelhos o número “1” e às pessoas com cabelos castanhos “2” como forma de medida. Neste caso, os números são atribuídos aos indivíduos de uma forma sistemática

de maneira que as diferenças entre os escores representam as diferenças na propriedade que está sendo medida (cor do cabelo). Da mesma forma, dar a um examinando uma pontuação 98 em uma prova de matemática ou a pontuação 54 num teste de personalidade seria a realização de uma medida se os números forem sistematicamente atribuídos para representar as diferenças de desempenho nos testes de matemática ou de personalidade". (p. 2) ¹

3.1.1. A medição nas Ciências Sociais

No entanto, quando o assunto é fazer medições, é comum ocorrer uma distinção entre os processos de medição nas ciências exatas (como a física e a química) e as demais ciências (como a psicologia, sociologia, economia, antropologia, ensino e aprendizagem em física, etc). Essa separação parece sugerir que as ciências como a Física constroem afirmações verdadeiras, enquanto as demais ciências não [Babbie 2005], sendo que esse ponto de vista tem sua origem na discussão do status "científico" das áreas até então englobadas sob a rubrica de "ciências sociais". O ponto central dessa separação reside no questionamento de se saber se o comportamento humano pode ou não ser submetido ao estudo "científico".

É importante se ressaltar que esta visão que coloca em dúvida o status científico das "ciências sociais", encontra voz [Babbie 2005] tanto no campo das ciências sociais, onde profissionais acostumados aos métodos tradicionais nessas áreas se opõem às novas orientações, quanto nas áreas das ciências exatas, que afirmam que o "método científico" não pode ser aplicado ao comportamento social humano.

O primeiro ponto a ser levado em consideração, na desconstrução da ideia que não é possível se realizar medições no campo das ciências sociais, é o fato de que as teorias, em qualquer área, quase sempre resultam de uma

¹ Original: "*Measurement* is the assigning of numbers to individuals in a systematic way as a means of representing properties of the individuals. Numbers are assigned to the individuals according to a carefully prescribed, repeatable procedure. For example, a personality test generates scores by using the same instructions, questions, and scoring procedures for each examinee. Scores couldn't be compared meaningfully if examinees were each given different instructions or items or if different scoring procedures were used. In measurement, numbers are assigned systematically and can be of various forms. For example, labeling people with red hair "1" and people with brown hair "2" is measurement, since numbers are assigned to individuals in a systematic way and differences between scores represent differences in the property being measured (hair color). Similarly, giving an examinee a score of 98 on a mathematics exam or a score of 54 on a personality test would be measurement, if the numbers were systematically assigned to represent differences in mathematics performance or personality." (p. 2)

combinação de longos processos dedutivos e indutivos [Babbie 2005], ou seja, tendo-se uma explicação inicial para uma relação empírica, esta pode ser testada e o resultado deste teste utilizado para modificar a ideia inicial que se tinha para essa explicação.

Como um segundo argumento para invalidar essa diferenciação, tem-se que quase nenhum experimento, seja na área das ciências exatas ou sociais, é crítico, ou seja, é capaz de derrubar ou confirmar de forma absoluta e incontestável uma teoria. Isto leva a se concluir que nenhum cientista, seja da área que for, será capaz de descobrir “A Verdade” como conhecimento absoluto e incontestável [Babbie 2005].

Como último argumento em defesa da desconstrução da ideia mencionada, tem-se que a operacionalização desses conceitos depende da interpretação de quem os estabelece e os utiliza. Além disso, a avaliação das hipóteses e teorias da qual elas tenham sido derivadas requer do cientista tomadas de decisões que indubitavelmente influenciarão o resultado do experimento, deixando claro que os resultados obtidos sempre possuirão um cunho pessoal dos que realizam a referida pesquisa.

Dentro desse contexto, portanto, observa-se que tanto a área de ciências sociais quanto a de ciências exatas apresentam problemas semelhantes na forma de se obter informações e construir suas teorias, indicando assim a possibilidade de realização de medidas em qualquer uma dessas áreas. Segundo Babbie, “*o comportamento social humano pode ser submetido a um estudo “científico” tão legitimamente quanto átomos e células*” [Babbie 2005] e este processo, percorrido tanto pelos cientistas físicos quanto pelos das ciências sociais, pode ser representado como na Figura 3.1.

Na busca de regularidades sociais, o primeiro passo a ser tomado num estudo científico é a realização de “medidas” de forma sistemática, seja de características, comportamentos ou atitudes sociais (aptidões)².

Como exemplo de características sociais que são passíveis de medição tem-se a cor de pele, idade, cidade natal, sexo, renda familiar [Babbie 2005],

² A definição de aptidão aqui adotada é a utilizada por Pasquali [Pasquali 2009] em seu livro *Psicometria: Teoria dos Testes na Psicologia e na Educação*, nas páginas de 55 a 61.

sendo que a especificidade e a quantidade das informações coletadas só dependerão da necessidade do pesquisador.

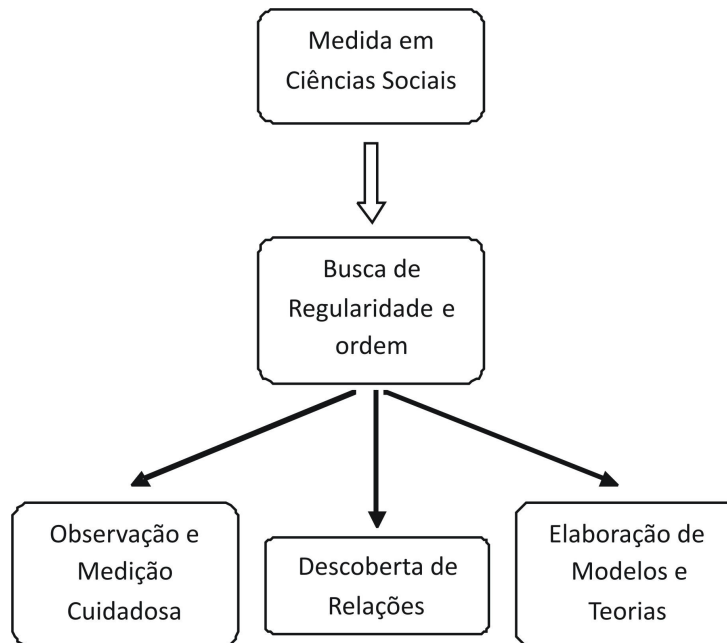


Figura 3.1. Aspectos da Obtenção de Medidas em Ciências Sociais

Em relação à medição dos comportamentos sociais, pode-se citar o resultado obtido nas urnas eleitorais, a medida das quantidades de tráfego em locais específicos ou o estudo do mercado de consumo.

Um dos pontos de conflito na aceitação de realizar medidas em “ciências sociais” é a medição de atitudes. No entanto, semelhante às ciências físicas, para se realizar essas medições, é possível criar conceitos operacionais de atitudes como o racismo, a religiosidade e a aprendizagem, entre outros. A medição destes conceitos, por sua vez, não se constitui da medida da atitude em si, porém, dentro dos parâmetros estabelecidos pelo cientista, fornecerá uma medida operacional que representará a atitude que se deseja medir. Segundo Babbie [Babbie 2005]

“Deve-se reconhecer que todas estas medidas (todas medidas, aliás) são basicamente arbitrárias. O cientista social não pode descrever uma pessoa inequivocamente como “alienada” e outra como “não alienada”. Pessoas serão, ao invés, descritas como relativamente mais ou menos alienadas – ou seja, comparando uma com a outra. Esta característica, entretanto, não é prerrogativa das ciências sociais, como demonstram a “escala de dureza” usada nas ciências físicas, a “escala Ritchter” para terremotos, etc. Ninguém pode dizer que um metal é “duro” ou que um terremoto é “severo”, apenas que é mais “duro” ou mais “severo” que o outro.” (p. 59)

Embora diferentes das existentes nas ciências exatas, há regularidades também nas ciências sociais. Isto pode ser observado, por exemplo, nas regularidades constatadas nas normas sociais, sejam estas prescritas por instituições formais da sociedade ou não.

A existência dessas regularidades, no entanto, é colocada em cheque através de argumentos relacionados à sua trivialidade, à existência de casos contraditórios e a possibilidade de perturbação da ordem pelos componentes que constituem o grupo de pesquisa que compõem as regularidades medidas [Babbie 2005].

Em relação à trivialidade das informações medidas, argumenta-se que no “senso comum” pode-se encontrar muitas generalizações falsas e informações equivocadas, sendo essencial que se teste as mesmas empiricamente, mesmo quando estas não são questionadas por leigos. Babbie [Babbie 2005] comenta que

“Muitos instrutores de metodologia das ciências sociais iniciam suas aulas revelando um conjunto de “descobertas importantes” das ciências sociais, derivadas de estudos conduzidos por Samuel A. Stouffer durante a II Guerra Mundial. Tais descobertas incluem os seguintes achados:

- *Soldados negros ficavam mais felizes em campos de treinamento no Norte do que no Sul dos Estados Unidos;*
- *Soldados da Força Aérea americana, onde as promoções eram rápidas, tinham mais probabilidade de achar que seu sistema de promoções era justo do que os soldados da Polícia Militar, onde elas eram muito lentas;*
- *Soldados mais escolarizados tinham mais probabilidade de se ressentir com o alistamento militar forçado que soldados menos escolarizados.*

Quando os alunos começam a fazer pouco caso da obviedade das “descobertas importantes”, o instrutor revela que cada uma delas foi desmentida pela pesquisa de Stouffer.” (p. 60)

Observa-se então que seja nas ciências exatas ou sociais, a documentação do “óbvio” é essencial para que se possua uma “interpretação correta” do fenômeno estudado, não se constituindo, então, como uma crítica legítima para se desencorajar qualquer pesquisa científica.

Em relação ao fato de se possuir exemplos de casos contraditórios nas pesquisas em ciências sociais, deve-se suscitar que, assim como em muitos casos das ciências exatas, as regularidades obtidas são padrões probabilísticos,

não necessitando que a relação entre duas variáveis seja verdadeira em 100% dos casos observáveis. Babbie afirma que [Babbie 2005]

“Na genética, por exemplo, o cruzamento de uma pessoa de olhos azuis com outra de olhos castanhos provavelmente resultará numa criança de olhos castanhos. Mas se a criança nascer de olhos azuis, isto não desafia a regularidade observada, já que a genética afirma apenas que a probabilidade dos olhos serem castanhos é maior e, mais ainda, que olhos castanhos devem ser esperados num certo percentual dos casos. O cientista social faz uma previsão probabilística semelhante – que, nos cálculos gerais, mulheres têm mais probabilidade de serem mais religiosas do que homens. Com instrumentos de medição testados adequadamente, podemos prever o percentual de mulheres que se mostrarão mais religiosas que os homens.” (p.61)

A possibilidade de haver perturbação nas regularidades sociais por vontade consciente dos indivíduos envolvidos não invalida o trabalho desenvolvido nessa área. Essas perturbações podem ocorrer, no entanto não com frequência suficiente para colocar em cheque a existência de regularidades sociais. Segundo Babbie [Babbie 2005]

“Isto não é negar que um racista branco de direita pode, se quiser perturbar os cientistas políticos estudando eleições, votar num negro esquerdista radical. Numa eleição, todos os votantes podem, de repente, mudar e votar no candidato azarão, frustrando as pesquisas eleitorais. Seguindo o mesmo raciocínio, todos os trabalhadores podem ir trabalhar mais cedo ou ficar em casa e não ir trabalhar, evitando que aconteça o congestionamento de tráfego esperado na hora do “rush”. (...) Ironicamente, é claro, se todos os trabalhadores ficarem em casa, isto também seria uma regularidade susceptível de explicação. (...) O fato é que normas sociais existem e podemos observá-las. Quando mudam com o tempo, podemos também observar e explicar mudanças. Em última análise, regularidades sociais persistem porque tendem a fazer sentido para os indivíduos nelas envolvidos.” (p. 61 e 62)

3.1.2. A Operacionalização de Conceitos Abstratos

As medições realizadas numa pesquisa possuem basicamente duas funções: a descritiva, quando se estuda como os resultados experimentais se distribuem nas variáveis estudadas, e a explicativa, obtida através da relação entre essas variáveis.

No entanto, no desenvolvimento de trabalhos de pesquisa é comum se deparar com a necessidade de se medir conceitos abstratos como classe social, racismo ou a aprendizagem em física.

Por serem abstratos, esses conceitos, em geral, não passam de ideias gerais na mente do pesquisador. Para que se possa medi-los, é necessário o levantamento de características desses conceitos para que se estabeleça uma definição operacional do mesmo. A partir dessa definição, então, reduz-se esses conceitos gerais a indicadores empíricos específicos que não representarão a definição completa e incontestável desse conceito, mas sim características do mesmo que o pesquisador julga úteis para compreensão de seus questionamentos.

Tem-se, então, que o estabelecimento desses conceitos apresenta uma multiplicidade de aspectos, de acordo com o objetivo e utilidade desejada pelo pesquisador. Por exemplo, ao se perguntar a professores que atuam no ensino médio em escolas brasileiras sobre o que entendem por aprendizagem em física, alguns afirmariam que saber física é saber resolver problemas numéricos de física; outros diriam que é solucionar problemas que necessitam de conhecimento dos conceitos físicos estudados, ou ainda, que essa aprendizagem se constituiria no aluno conseguir aplicar os conceitos estudados em seu cotidiano. Cada uma dessas posições apresenta uma característica que pode ser utilizada para uma definição operacional do que seria aprendizagem em física.

Já para o PISA (Programme for International Student Assessment), a aprendizagem em ciências, área em que a física está compreendida, significa adquirir “letramento científico”, ou seja, possuir a capacidade de usar o conhecimento científico para identificar questões e chegar a conclusões baseadas em evidências para entender e ajudar a tomar decisões a respeito do mundo e as mudanças causadas a ele pela atividade humana [OECD 2000].

Os Parâmetros Curriculares Nacionais para o Ensino Médio, o PCN+, em suas orientações curriculares para o Ensino de Física, definem como aprendizagem em física o resgate do espírito questionador do aluno e seu desejo de explorar o mundo, reconhecendo a física como cultura e como possibilidade de conhecer este mundo que o cerca. Neste processo, o aluno deve desenvolver uma atitude reflexiva e autocrítica diante dos erros cometidos, gerenciar os

conhecimentos adquiridos e compreender a predominância de aspectos técnicos e científicos na tomada de decisões sociais significativas e os conflitos gerados, nestes, pela negociação política [PCN+ 2002].

Como um último exemplo, no Exame Nacional do Ensino Médio (Novo ENEM), aprender física significa que o aluno adquiriu ou desenvolveu um conjunto de habilidades pré-determinadas que, em conjunto, representam a aquisição e o desenvolvimento de algumas competências (aptidões) elencadas na Matriz de Referência desse exame.

Então, como se pode perceber, conceitos tidos como abstratos dependem da definição operacional atribuída a eles, não apresentando dessa forma um significado último e universal. Logo, não faz sentido dizer que o conceito operacional de um conceito abstrato está certo ou errado, mas sim classificá-lo como mais ou menos útil para responder às questões propostas.

É necessário refletir sobre as consequências desta discussão. A partir do momento que o pesquisador define o conceito operacional e quais aspectos desse conceito serão estudados, pode-se dizer que ele não coleta dados para o seu estudo, mas sim que cria esses dados.

Então, pode-se concluir que medir um conceito abstrato, muitas vezes, implica em “recriar” este conceito através de parâmetros concretos que sejam possíveis de ser medidos, gerando assim um novo conceito (o conceito operacional), que apesar de útil, não é a medida real do conceito abstrato em questão. Esquemáticamente, tem-se:

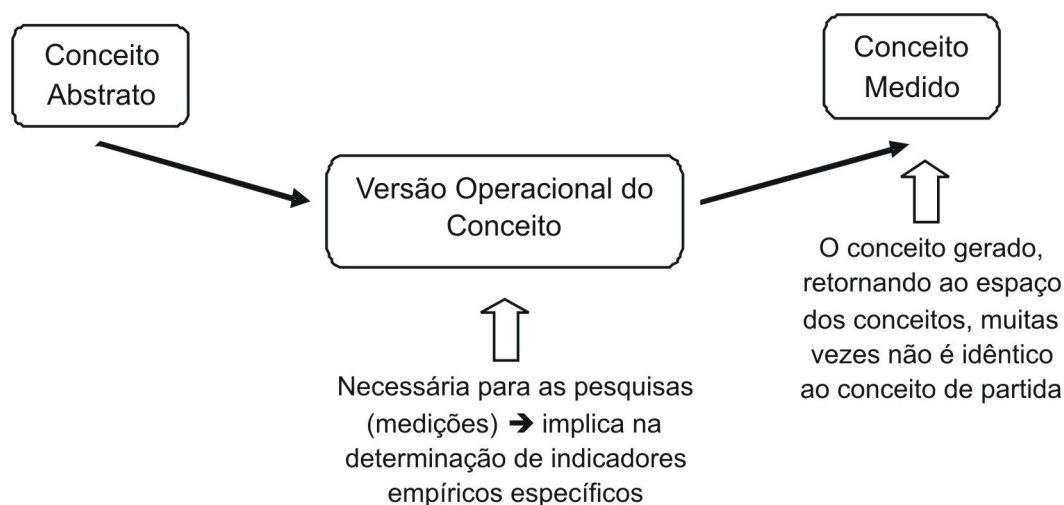


Figura 3.2. Aspectos da Obtenção de Medidas em Ciências Sociais

Definido o conceito operacional referente a um conceito abstrato, para que se proceda às medidas necessárias, é preciso elaborar uma extensa e detalhada lista das variáveis que podem ser utilizadas na construção dessa definição, explicitando-se também aí os atributos constituintes dessas variáveis. Por exemplo, quando se faz o estudo do rendimento em física em um grupo de estudantes, pode-se propor a resolução de alguns exercícios; e pode-se pesquisar se a variável “gênero”, que possui os atributos “masculino” e “feminino”, interfere nesse resultado. Se quer-se investigar se o rendimento pode estar relacionado ao nível de formação dos pais, cria-se mais uma variável, que pode possuir como atributos “ensino fundamental”, “ensino médio”, “graduação” e “pós-graduação”.

Observa-se, que esses atributos também podem ser representados ao longo de uma dimensão. Por exemplo, ao se relacionar o rendimento do grupo de estudantes, acima citados, ao nível de compromisso na realização das tarefas determinadas pelo professor, pode-se indicar como atributos dessa variável o “baixo compromisso”, o “médio compromisso” e o “alto compromisso”.

Ressalte-se que junto à listagem das variáveis utilizadas para caracterizar o conceito definido, deve-se também listar o que não é ou o que não caracteriza este conceito. Esse procedimento permite a construção de uma gradação entre o que é o conceito definido e o que não é, enriquecendo assim as informações da pesquisa realizada. Por exemplo, ao se definir o que é aprendizagem em física, deve-se definir o que não é aprendizagem, estabelecendo-se aí dois extremos do conceito aprendizagem em física, podendo-se então buscar uma gradação entre esses conceitos.

Finalizando essa discussão, pode-se definir como a “operacionalização” de um conceito [Babbie 2005] como o processo pelo qual são especificadas observações empíricas que podem ser tomadas como indicadores dos atributos contidos deste conceito.

3.2- Exemplos de Medidas

Pode-se agora exemplificar como um mesmo conceito pode ser tratado ou medido de diferentes formas, fornecendo diferentes visões sobre este conceito.

3.2.1. Aprendizagem de Conceitos Relativos ao Movimento de Corpos Próximos à Superfície da Terra – a Aceleração da Gravidade.

Fazendo parte da mecânica, conteúdo abordado na maior parte dos currículos das escolas do ensino médio, os movimentos de projéteis próximos a superfície da Terra (queda livre e lançamento vertical no vácuo) são ensinados como um caso particular dos movimentos retilíneos uniformemente variados (MRUV), cuja principal característica é possuírem aceleração constante.

No ensino tradicional, que constitui a forma de trabalho da maior parte das escolas de ensino médio, o estudo deste tema compreende trabalhar exaustivamente com equações e gráficos na resolução de problemas numéricos. Coerente com esse trabalho, a avaliação da aprendizagem desse conteúdo é feita através de provas e trabalhos que contemplam os mesmos tipos de exercícios numéricos.

Um exemplo deste tipo de questão, muito utilizada, pode ser observado no exemplo a seguir.

Um corpo é lançado verticalmente para cima, em um local onde o efeito do atrito com o ar é desprezível. Se a velocidade de lançamento foi de 30 m/s qual a altura máxima atingida pelo corpo? (Considere $g = 10 \text{ m/s}^2$)

Pode-se utilizar a equação de Torricelli, $V^2 = V_0^2 + 2.a.\Delta s$ para a resolução deste problema. Nesta equação, “V” é a velocidade final, “V₀” a velocidade inicial, “a” a aceleração e “Δs” o deslocamento realizado pelo móvel.

Levando-se em consideração que ao atingir a altura máxima a velocidade do corpo se anula, os estudantes procedem normalmente com a resolução da questão, conforme ilustrado abaixo:

$$\begin{aligned}V^2 &= V_0^2 + 2.a.\Delta s \\0^2 &= 30^2 + 2.(-10).\Delta s \\20\Delta s &= 900 \\ \Delta s &= 45m\end{aligned}$$

O que se observa quando se corrige essa questão é que o aluno, para o ponto de altura máxima e conseqüentemente de velocidade instantânea zero, substitui o valor da aceleração da gravidade por 10 m/s^2 , indicando que ele sabe

que a aceleração da gravidade não se anula no ponto de altura máxima. No entanto, quando questionado conceitualmente sobre este mesmo ponto, há indicações que o mesmo não compreendeu o conceito inquirido.

Para um professor que utiliza somente este tipo de problemas, medir a aprendizagem em física nesse conteúdo consiste na memorização e aplicação de equações para resolver problemas numéricos. Observe que, dentro dos parâmetros estabelecidos pelo professor para o conceito de aprendizagem deste conteúdo, existe coerência em suas ações e ele pode afirmar que o aluno que acerta o exercício aprendeu o conteúdo ministrado.

No entanto, se a concepção de aprendizagem dos conceitos físicos deste conteúdo implicar no conhecimento e compreensão das grandezas físicas envolvidas no fenômeno, observa-se que grande parte dos alunos que resolvem o problema numérico do exemplo anterior não percebe a gravidade como grandeza física, não identifica suas características e tão pouco possui consciência da ação dela sobre os corpos.

Observe as questões abaixo, apresentadas nas Figuras 3.3 e 3.4. Elas foram aplicadas a alunos calouros de um curso da área de ciências exatas da UFRJ, que na sua maioria são ingressam vindos de escolas tradicionais.

12) Uma bola é lançada verticalmente para cima. No ponto mais alto da trajetória da bola,

Resposta	Média	Total
sua aceleração é nula, e sua velocidade é não nula.	8.5%	8
sua velocidade e aceleração são nulas.	57.4%	54
sua velocidade é nula, mas a aceleração não é nula.	35.1%	33

Figura 3.3. Questão 12, aplicada aos estudantes calouros da UFRJ do curso de Ciências Matemáticas e da Terra em 2010 [Estudantes calouros UFRJ, comunicação privada, MFB]

Como se pode notar, mais da metade dos estudantes acredita que a aceleração no ponto mais alto da trajetória é nula, ou seja, a execução de exercícios numéricos de física, feita pela maior parte dos alunos no ensino médio, não garantiu a compreensão do conceito da grandeza aceleração da gravidade para esta maioria. Esse fato é reforçado quando observa-se a próxima questão, aplicada ao mesmo grupo de alunos, na Figura 3.4.

13) Considere duas situações:

Situação I - Uma bola é lançada para cima, verticalmente.

Situação II - Uma bola é largada do alto de uma torre.

Podemos afirmar que

Resposta	Média	Total
a aceleração depende da velocidade com que a bola é lançada na situação I, e da altura que é largada na situação II.	24.5%	23
na primeira situação, a aceleração é vertical e para cima, e na segunda a aceleração é vertical e para baixo.	51.1%	48
nada podemos afirmar sobre a aceleração, pois não temos nenhuma informação sobre as velocidades.	1.1%	1
nas duas situações, a bola tem a mesma aceleração.	11.7%	11
Nenhuma das respostas anteriores está correta.	12.8%	12

Figura 3.4. Questão 13, aplicada aos estudantes calouros da UFRJ do curso de Ciências Matemáticas e da Terra em 2010 [Estudantes calouros UFRJ, comunicação privada, MFB].

No entanto, se a concepção de aprendizagem dos conceitos físicos deste conteúdo implicar no conhecimento e compreensão das grandezas físicas envolvidas no fenômeno, observa-se que grande parte dos alunos que resolvem o problema numérico do exemplo anterior não percebe a aceleração da gravidade como grandeza física, não identifica suas características e tão pouco possui consciência da ação dela sobre os corpos.

Esses resultados apontam para o fato que saber resolver problemas numéricos não garante a aprendizagem do conceito aceleração da gravidade. Em outras palavras, um professor que entende a compreensão do conceito como aprendizagem do mesmo não terá nos problemas numéricos bons parâmetros para sua avaliação.

Pesquisas em Ensino de Física indicam [Arons 1997] que existe uma dificuldade tremenda dos estudantes em situações como a contemplada nas duas questões respondidas pelos calouros da UFRJ. Logo, o conceito de aprendizagem em física adotado pelo professor exemplificado no primeiro problema (numérico) não é útil quando se deseja que o estudante adquira uma compreensão maior sobre o conceito de aceleração da gravidade, pois, apesar de utilizar matematicamente o valor desta aceleração para o cálculo da altura atingida pelo

corpo no primeiro problema apresentado, na maior parte das vezes ele não consegue relacionar a operação matemática com o conceito solicitado nas duas questões.

Isto é uma evidência do fato de que, ao operacionalizarmos um conceito abstrato, podemos ter uma medida mais útil ou menos útil para o conceito, de acordo com o que deseja saber o avaliador.

3.2.2. Mudança de Fase – Existe Água a Zero Graus?

Muitos alunos associam a fusão do gelo ou a solidificação da água exclusivamente à temperatura de fusão, ignorando que, ao atingir esta temperatura, é necessária a continuidade do fluxo de calor, seja ele cedido ou recebido, para que o fenômeno ocorra [Gonçalves Junior e Barroso 2011a].

No sistema tradicional de ensino, hegemônico nas escolas brasileiras, esse assunto é abordado dando ênfase a problemas numéricos que implicam no cálculo da quantidade de calor cedida ou recebida para que uma substância mude de temperatura ou estado de agregação da matéria.

Com o objetivo de ilustrar esse tipo de problema, suponha que se deseja calcular a quantidade de calor absorvida por 100 gramas de gelo a -10°C para ele se transforme em 100 gramas de água a 20°C . Para isso se deve considerar o calor latente de fusão do gelo $L_{\text{fusão}} = 80 \text{ cal/g}$, calor específico da água $c_{\text{água}} = 1,0 \text{ cal/g}^{\circ}\text{C}$, calor específico do gelo $c_{\text{gelo}} = 0,5 \text{ cal/g}^{\circ}\text{C}$ e a temperatura de fusão do gelo 0°C .

Como resolução do problema, tem-se:

O gelo a -10°C tem que atingir sua temperatura de fusão (0°C), para então começar mudar de fase; logo, inicialmente tem-se que calcular a quantidade de calor absorvido nesse processo, o que é feito através da relação $\Delta Q = m.c.\Delta\theta$.

Ao atingir a temperatura de 0°C , o gelo começa fundir-se e a temperatura mantém-se constante até que todo gelo tenha se transformado em água. Neste processo o calor absorvido é calculado pela relação $\Delta Q = m.L_{\text{fusão}}$.

Enfim, quando todo gelo se transformar em água, a temperatura do líquido volta a subir até atingir 20°C . Novamente, neste processo, o calor absorvido é calculado por $\Delta Q = m.c.\Delta\theta$.

Na Figura 3.5, apresenta-se um esquema desta solução.

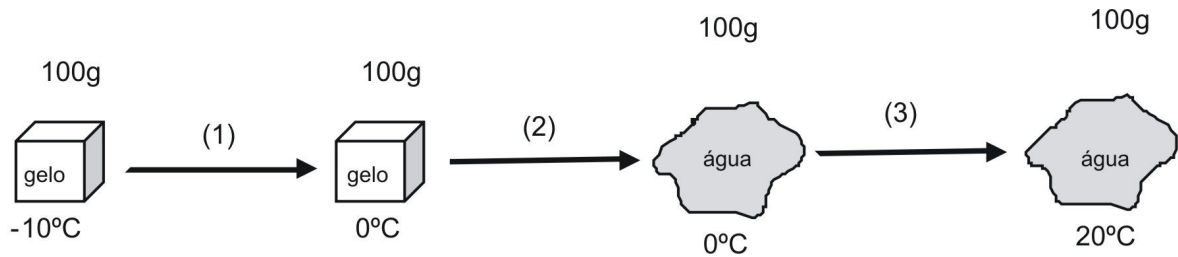


Figura 3.5. Esquema representando a transição do gelo a -10°C a água a 20°C .

Mais especificamente, em cada etapa:

(1) não houve mudança de fase \rightarrow calor sensível $\rightarrow \Delta Q = m.c.\Delta\theta \rightarrow \Delta Q = 100 \cdot 0,5 \cdot [0 - (-10)] \rightarrow \Delta Q_{(1)} = 500$ calorias

(2) houve mudança de fase \rightarrow calor latente $\rightarrow \Delta Q = m.L_{\text{fusão}} \rightarrow \Delta Q = 100 \cdot 80 \rightarrow \Delta Q_{(2)} = 8000$ calorias

(3) não houve mudança de fase \rightarrow calor sensível $\rightarrow \Delta Q = m.c.\Delta\theta \rightarrow \Delta Q = 100 \cdot 1 \cdot (20 - 0) \rightarrow \Delta Q_{(3)} = 2000$ calorias

Logo o calor total absorvido pelos 100 gramas de gelo inicial é $\Delta Q_{\text{total}} = \Delta Q_{(1)} + \Delta Q_{(2)} + \Delta Q_{(3)} \rightarrow \Delta Q_{\text{total}} = 10.500$ calorias

Após a resolução desse tipo de problema, o professor que utiliza este tipo de exercício como parâmetro para determinar a aprendizagem do aluno pode inferir que o estudante que a fez corretamente aprendeu os conceitos relacionados à troca de calor e mudança de estado de agregação da matéria.

No entanto, esse parâmetro se torna menos útil quando colocado o seguinte problema para avaliar a aprendizagem destes mesmos conceitos:

Suponha a existência de dois freezers que mantenham a temperatura dentro deles constantemente a 0°C . No primeiro, coloca-se gelo a uma temperatura inicial de -10°C e, no segundo, coloca-se água a 20°C . Se os freezers forem reabertos após alguns dias, o que você espera encontrar nas vasilhas que inicialmente continham gelo e água?

A resposta correta é que se encontrará gelo onde havia gelo e água onde havia água. No entanto é considerável o número de alunos que afirmam que encontrariam gelo nas duas vasilhas. Esta resposta indica que o aluno associa a

mudança de fase à temperatura de solidificação da água e não à necessidade de continuar perdendo calor após atingi-la, isto é, ele acredita que não pode existir água a 0°C [Gonçalves Junior e Barroso 2011a].

Observa-se, novamente, que é necessário ter-se muito claro qual é o conceito operacional do conceito abstrato a ser medido, a fim de que se possa fazer a escolha de parâmetros que sejam úteis para o estudo pretendido.

3.2.3. Conclusão

Como se pode concluir ao final dessa discussão, a medição de conceitos abstratos está longe de ser algo simples ou fácil. A impossibilidade de medi-los diretamente obriga ao pesquisador criar um processo de operacionalização próprio, que nem sempre será aceito e compartilhado por todos seus pares.

É fundamental, então, nesse processo, que se compreenda que essa operacionalização pode ser feita de diversas maneiras e sempre será parametrizada e direcionada pelo seu criador, não devendo por isso ser classificada como certa ou errada, mas simplesmente, de acordo com seus critérios, ela será mais ou menos útil quem por ventura venha utilizá-la.

Por isso, ao se pensar em medir um conceito abstrato, deve-se buscar uma definição operacional que satisfaça os questionamentos do pesquisador, possibilitando assim produzir análises a partir deles.

3.3- As técnicas de avaliação de larga escala: TCT e TRI

Nos processos de avaliação de aprendizagem, precisa-se atribuir um escore (uma nota), em geral comparativa, aos estudantes. O que usualmente é feito é atribuir uma pontuação a uma série de questões ou problemas, e somar a pontuação obtida pelo aluno para obter uma nota final. Este mecanismo de obtenção do escore a ser atribuído ao aluno é característica da denominada Teoria Clássica dos Testes (TCT).

Alguns dos problemas enfrentados na Teoria Clássica dos Testes (TCT), quando se pensa em avaliações de larga escala, é a viabilidade de comparação

entre os grupos respondentes e a dependência existente entre as características do teste e do examinando. É conhecido de todos os professores que existem provas ou testes “fáceis” ou “difíceis”, turmas “fracas” e “fortes”. Para construir uma forma de elaboração de testes que permita comparações com significados entre grupos e épocas diferentes, utiliza-se a denominada Teoria da Resposta ao Item, que transpõe algumas das limitações observadas na Teoria Clássica de Testes.

3.3.1. A Teoria Clássica de Testes (TCT)

Na Teoria Clássica de Testes o foco está na produção de testes de qualidade que no final resultarem em testes válidos. Isso se deve ao fato de que ela se ocupa da explicação do resultado final total do mesmo, ou seja, o escore total que é dado pela soma das respostas corretas de um dado conjunto de itens [Pasquali 2003].

Por exemplo, no questionário de física térmica discutido a seguir, no Capítulo 4, tem-se 26 questões a serem respondidas pelos alunos. Um determinado aluno (por exemplo o de número 46), ao respondê-lo, acertou 14 itens. Como foi atribuído grau 0 para cada erro e grau 1 para cada um dos acertos, pode-se afirmar que este aluno, neste teste, apresentou um escore total igual a 14. A TCT, então, se ocupa em dar significado ao que representa esse escore para o referido aluno³.

No entanto, quando se deseja comparar grupos diferentes em momentos diferentes utilizando testes diferentes, a TCT apresenta limitações, das quais algumas são enumeradas a seguir [Hambleton 1991, Pasqualli 2003]:

1) Os parâmetros clássicos utilizados para analisar um item (uma questão) de um teste (uma prova), a dificuldade do item e a capacidade de discriminação do item, dependem diretamente da amostra de sujeitos utilizada

³ É interessante observar que a aplicação de testes comportamentais [Pasquali 2009] se popularizou durante a Primeira Guerra Mundial, quando o exército americano passa a utilizá-los, individual e coletivamente, para selecionar seus soldados. Logo após esse período, diversas indústrias e instituições passam também a adotá-los na seleção de seus empregados. Paralelamente, são desenvolvidos e utilizados testes e inventários de personalidade, além de instrumentos menos objetivos denominados de testes projetivos.

para estabelecê-los (group-dependent ou sample-dependent). Se a amostra não for rigorosamente representativa da população, os parâmetros dos itens não podem ser considerados válidos para esta população. A dificuldade de se obter amostras representativas é um problema prático grave para os construtores de testes, e a dependência dos parâmetros dos itens com a amostra torna-se um obstáculo quase intransponível quando se deseja elaborar instrumentos psicométricos não tendenciosos.

Por exemplo, quando um professor deseja determinar o nível de dificuldade de um teste feito por ele, normalmente ele pede a um colega de sua área para avaliar este teste ou então verifica o resultado do mesmo numa turma que já o tenha feito. No entanto, ao aplicá-lo em outros grupos de alunos, ele percebe que o teste apresenta níveis de dificuldade diferentes, de acordo com as características de cada grupo, ou seja, o resultado da amostra (dos alunos que fizeram o testes) não é válido para análise das outras turmas que não fizeram o teste. Esta característica da TCT extrapola os muros das instituições de ensino, ocorrendo, por exemplo, nas pesquisas de opinião.

2) O teste que é aplicado interfere diretamente na avaliação das aptidões dos estudantes que realizam este teste (test-dependent). Ao se aplicar testes diferentes para medir a mesma aptidão, obtém-se escores diferentes da mesma aptidão para sujeitos idênticos. Os escores também serão diferentes quando se aplica testes de dificuldades diferentes. No caso das formas paralelas de testes, é preciso observar que, em primeiro lugar, conseguir formas estritamente paralelas é uma tarefa quase impossível e, em segundo lugar, mesmo conseguindo formas paralelas, é difícil pressupor que elas produzam o mesmo montante de erro, o que vem afetar a estimação do escore verdadeiro dos sujeitos.

Observe, por exemplo, os dois itens apresentados a seguir, na Figura 3.6, constantes do vestibular da UFRJ, respectivamente nos anos de 1999 e 2004.

Os dois itens compreendem a aprendizagem de conceitos relacionados à decomposição de grandezas vetoriais e o equilíbrio da partícula. No entanto, no item de 2004, a colocação da força \vec{F} numa direção não comumente utilizada nas resoluções de planos inclinados torna o item mais difícil que o de 1999. Supondo, então, que essas duas questões fossem aplicadas para determinar a aptidão dos alunos em relação a estes conceitos, o fato de a segunda questão ser mais difícil

faria com que um mesmo aluno tivesse valores diferentes para sua aptidão dependendo de qual dos dois itens lhe fosse proposto.

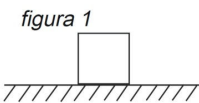
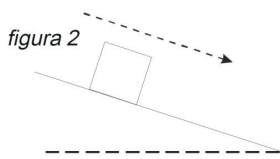
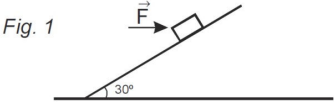
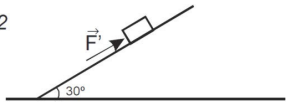
<p>QUESTÃO 2</p> <p>A figura 1 mostra um bloco em repouso sobre uma superfície plana e horizontal. Nesse caso, a superfície exerce sobre o bloco uma força \vec{f}. A figura 2 mostra o mesmo bloco deslizando, com movimento uniforme, descendo uma rampa inclinada em relação à horizontal segundo a reta de maior declive. Nesse caso a rampa exerce sobre o bloco uma força \vec{f}'.</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>figura 1</p> </div> <div style="text-align: center;">  <p>figura 2</p> </div> </div> <p>Compare \vec{f} e \vec{f}' e verifique se $\vec{f}' < \vec{f}$, $\vec{f}' = \vec{f}$ ou $\vec{f}' > \vec{f}$. Justifique sua resposta.</p>	<p>QUESTÃO 2</p> <p>Deseja-se manter um bloco em repouso sobre um plano inclinado 30° com a horizontal. Para isso, como os atritos entre o bloco e o plano inclinado são desprezíveis, é necessário aplicar sobre o bloco uma força. Numa primeira experiência, mantém-se o bloco em repouso aplicando uma força horizontal \vec{F}, cujo sentido está indicado na figura 1.</p> <div style="text-align: center;">  <p>Fig. 1</p> </div> <p>Numa segunda experiência, mantém-se o bloco em repouso aplicando uma força \vec{F}' paralela ao plano inclinado, cujo sentido está indicado na figura 2.</p> <div style="text-align: center;">  <p>Fig. 2</p> </div> <p>Calcule a razão \vec{F}' / \vec{F}.</p>
---	---

Fig. 3.6. Itens do vestibular da UFRJ 1999 (à esquerda) e 2004 (à direita).

3) Uma outra fonte de dificuldade da Teoria Clássica dos Testes está no conceito de precisão. Este conceito é definido na TCT como sendo a correlação entre os escores obtidos em formas paralelas de um teste (testes que apresentem mesmo nível de dificuldade e discriminação). O grande problema é que se obter formas paralelas de testes é praticamente impossível. Além disso, na TCT supõe-se que os desvios padrões dos escores obtidos são os mesmos para todos os alunos, o que é improvável quando se está analisando estudantes com aptidões diferentes.

Em outras palavras, construir dois testes que apresentem exatamente as mesmas características, avaliem as mesmas aptidões e ainda mais, apresentem a mesma dificuldade para todos os estudantes, afim de que sejam comparáveis, não é uma tarefa possível dentro desta teoria.

4) Outro problema da teoria clássica dos testes consiste em que ela é orientada para o teste total e não para o item individual. Toda a informação do item deriva de considerações do teste geral, não se podendo assim determinar como o examinando se comportaria diante de cada item individual.

Dentro dessas limitações, ao se aplicar um teste, a característica do examinando na qual se está interessado é a “aptidão” medida pelo instrumento de avaliação. A TCT expressa essa “aptidão” pelo escore verdadeiro que é definido como o valor esperado de desempenho observado no teste de interesse [Hambleton 1991]. Porém, na determinação desta aptidão na TCT enfrenta-se um grande problema que é o fato de que as características do examinando e do teste não podem ser separadas, ou seja, a aptidão do examinando é definida em termos de um teste particular. Isto é evidenciado quando se aplica um teste considerado difícil ao examinando, que neste caso pode obter um resultado que indique que o ele apresenta uma aptidão baixa. No entanto, se for aplicado outro teste que exija a mesma aptidão, mas este teste possuir um nível de dificuldade menor, pode-se obter, para o mesmo estudante, uma aptidão alta.

Tem-se, então, que na TCT as características do teste e do item mudam quando o contexto do examinando muda, e as características do examinando mudam quando o contexto do teste muda. Portanto, é muito difícil comparar examinandos que fazem testes diferentes, é muito difícil comparar itens cujas características são obtidas usando grupos diferentes de examinandos. Esse problema é constante na atuação dos professores que precisam fazer avaliações diferentes para turmas que estão cursando a mesma série, avaliações substitutas ou de segunda chamada. Por exemplo, um aluno que faz uma prova na data prevista e outro que por algum motivo necessita fazer a segunda chamada da mesma, mesmo tirando a mesma nota, não terão a mesma aptidão.

3.3.2. A Teoria de Resposta ao Item (TRI)

Buscando-se superar as limitações existentes na TCT, outras metodologias de avaliação foram propostas e, dentre elas, a Teoria de Resposta ao Item (TRI). Como características inerentes a esta teoria tem-se que [Hambleton, 1991]:

- “ 1) os itens (questões) não são dependentes do grupo que faz o teste.*
- 2) os escores descrevem a proficiência que independe do teste aplicado.*
- 3) esse modelo de avaliação é expresso a partir do item e não do teste como um todo, permitindo que se analise cada item individualmente, independentemente dos demais itens do teste.*

- 4) não há exigência de testes estritamente paralelos para a verificação de sua confiabilidade.
- 5) esse modelo fornece uma medida de precisão para cada escore de habilidade.” (p. 5)⁴

Nesse modelo de análise, busca-se medir o que se denomina por traço-latente ou variável não-observável. Para uma melhor compreensão do que isso significa, suponha que um professor de física, do ensino médio ou superior, ao avaliar o quanto seus alunos aprenderam (o que se constitui de um conceito abstrato), utiliza como instrumento para obter essa informação uma prova que contém alguns problemas; neste caso, a variável observada por ele é a capacidade de cada estudante resolver os problemas apresentados ou determinado tipo de problema e não há dúvidas que aprender algum tópico de física não é equivalente a saber resolver problemas daquele tópico.

Tabela 3.1. Habilidades (variáveis observáveis) pertencentes à competência 6.

Variáveis observáveis ou habilidades referentes à competência 6
H20 – Caracterizar causas ou efeitos dos movimentos de partículas, substâncias, objetos ou corpos celestes.
H21 – Utilizar leis físicas e (ou) químicas para interpretar processos naturais ou tecnológicos inseridos no contexto da termodinâmica e (ou) do eletromagnetismo.
H22 – Compreender fenômenos decorrentes da interação entre a radiação e a matéria em suas manifestações em processos naturais ou tecnológicos, ou em suas implicações biológicas, sociais, econômicas ou ambientais.
H23 – Avaliar possibilidades de geração, uso ou transformação de energia em ambientes específicos, considerando implicações éticas, ambientais, sociais e/ou econômicas.

Outro exemplo de medida de traço latente ou variável não observável pode ser encontrado na Matriz de Referência do ENEM⁵. Suponha que, ao se construir a prova de Ciências da Natureza, deseja-se medir a competência de se apropriar de conhecimentos da física para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas (competência 6), que

⁴ Original: “(a) item characteristics that are not group-dependent, (b) scores describing examinee proficiency that are not test-dependent, (c) a model that is expressed at the item level rather than at the test level, (d) a model that does not require strictly parallel test for assessing reliability, and (e) a model that provides a measure of precision for each ability score.” (p. 5)

⁵ Neste caso se utiliza o termo competência como sinônimo de traço latente (aptidão) e o termo habilidade como sinônimo de comportamento ou variável observável.

nesse documento representa uma variável não-observável ou traço latente. Para isso, constróem-se itens que contemplem as habilidades (variável observável) listadas na Tabela 3.1 que, em conjunto, possibilitam determinar essa competência (traço latente).

A Teoria da Resposta ao Item corresponde então a uma proposta psicométrica para determinação de aptidões (traços latentes) que possui dois postulados básicos:

Primeiro Postulado

O desempenho de um examinando num item do teste pode ser previsto (ou explicado) por um conjunto de fatores denominados de traços, traços latentes, ou aptidão.

Segundo Postulado

A relação entre o desempenho em um item e o conjunto de traços que definem este desempenho no item pode ser descrita por uma função monótonamente crescente da aptidão chamada função característica do item, ou curva característica do item (ICC=item characteristic curve).

Observe que o primeiro postulado traz consigo a idéia básica de que, através da realização de testes, é possível medir a aptidão do examinando.

O segundo postulado, por sua vez, apresenta uma idéia bastante razoável que consiste no fato de que ao se construir uma curva da probabilidade de acerto de um item para um aluno com determinada aptidão, essa curva deve assumir a forma de um S deitado, isto é, se este aluno possuir uma nota alta, a probabilidade de ele acertar qualquer item é também alta, se este aluno possuir nota média, a probabilidade de acertar qualquer item é média e finalmente, se o aluno tem nota baixa, a probabilidade também será baixa.

Considere, por exemplo, que um mesmo teste seja aplicado a duas turmas distintas que possuam níveis de aptidão diferentes, isto é, uma turma com menor nível de aptidão (turma 1) e a outra com maior nível de aptidão (turma2). Podemos construir uma curva que represente a distribuição dos escores do teste para as duas turmas. Certamente a turma 2 apresentará um valor médio mais

alto do que a turma 1. Na Figura 3.7, apresenta-se um gráfico onde a abscissa corresponde ao escore do aluno, ou melhor, à habilidade ou aptidão do aluno. Na parte inferior do gráfico, estão representadas (de forma invertida) as duas distribuições de notas para cada uma das duas turmas. Na parte superior da curva, representa-se a ICC dos dois grupos, isto é, a curva que representa, para um item do teste, a probabilidade de acerto no item para o candidato cujo escore é dado.

Observe que a ICC para as duas turmas é a mesma, sendo que se pode observar que os respondentes com maior aptidão (localizados mais à direita no eixo das abscissas) possuem também maior probabilidade de responder o item corretamente.

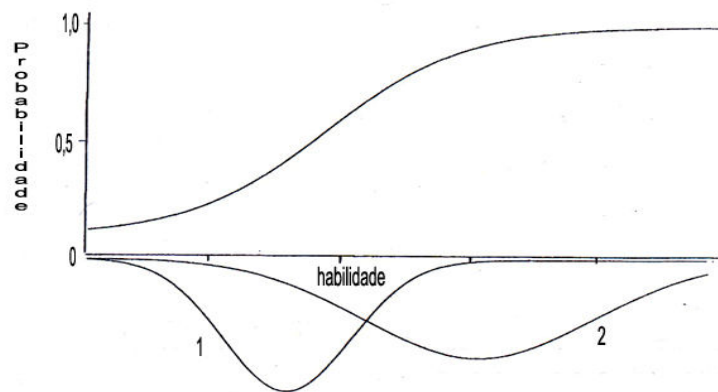


Fig. 3.7. ICC e curvas de distribuição de aptidão para as turmas 1 e 2 de examinandos [Hambleton 1991].

Como na utilização da metodologia da TRI os itens não são dependentes do grupo que é submetido ao teste, os escores que descrevem a aptidão de cada respondente independem do teste aplicado, e os parâmetros estimados em diferentes grupos também devem ser os mesmos, tanto o item quanto os parâmetros de aptidão são ditos invariantes [Hambleton 1991]. Isto pode ser observado na figura 3.7, onde dois examinandos que possuem a mesma aptidão, seja na turma 1 ou na turma 2, possuirão a mesma probabilidade de acertar um item.

Os modelos matemáticos que fundamentam a TRI especificam que a probabilidade de um examinando marcar a alternativa correta depende de sua aptidão ou do conjunto formado por esta aptidão e as características do item. A aplicação destes modelos estão alicerçados em três hipóteses:

1. A hipótese de unidimensionalidade do item: uma única aptidão é medida por cada item que compõe o teste. Esta característica é muito difícil de ser verificada, sendo que normalmente basta que exista uma componente ou fator dominante da referida aptidão que influencie no processo.
2. A independência local, que ocorre quando as aptidões que influenciam no desempenho no teste são mantidas constantes e as respostas a qualquer par de itens são estatisticamente independentes (ou seja, os únicos fatores que influenciam as respostas aos itens do teste são as aptidões especificadas no modelo)
3. E, finalmente, o fato que a função característica do item reflete a relação verdadeira entre as variáveis não observáveis (aptidões) e as variáveis observáveis (respostas aos itens).

3.3.3. Modelo de Rasch para itens dicotômicos (Compreendendo a idéia de uma ICC)

Para entender um pouco melhor as ideias gerais apresentadas anteriormente, apresenta-se um exemplo⁶ para discutir a construção de uma curva característica de um item.

Para examinar o desempenho – ou a “aptidão” – de um atleta de salto em altura, pode-se pensar em algumas alternativas para atribuir valores a ela:

- através de um recorde individual;
- através de um recorde individual durante um evento oficial ou internacional;
- através da média do desempenho do atleta durante um determinado período de tempo; ou
- através do desempenho mais freqüente num determinado período de tempo.

⁶ Exemplo retirado de PISA 2009 Data Analysis Manual, OECD.

http://www.oecd.org/document/19/0,3746,en_2649_35845621_48577747_1_1_1_1,00.html

A escolha do critério a ser usado para a medida desse desempenho impacta, por exemplo, na escolha de qual atleta vai participar de uma determinada competição.

Suponha, então, que para avaliar o desempenho, consideraram-se os resultados obtidos por dois atletas de salto em altura durante um ano. Constrói-se, como mostrado na Figura 3.8, um gráfico esquemático representando a probabilidade de sucesso desses atletas em função da altura saltada. Isso pode ser medido, por exemplo, fazendo o atleta repetir o salto numa altura dada várias vezes, e determinando o percentual de sucesso naquela altura.

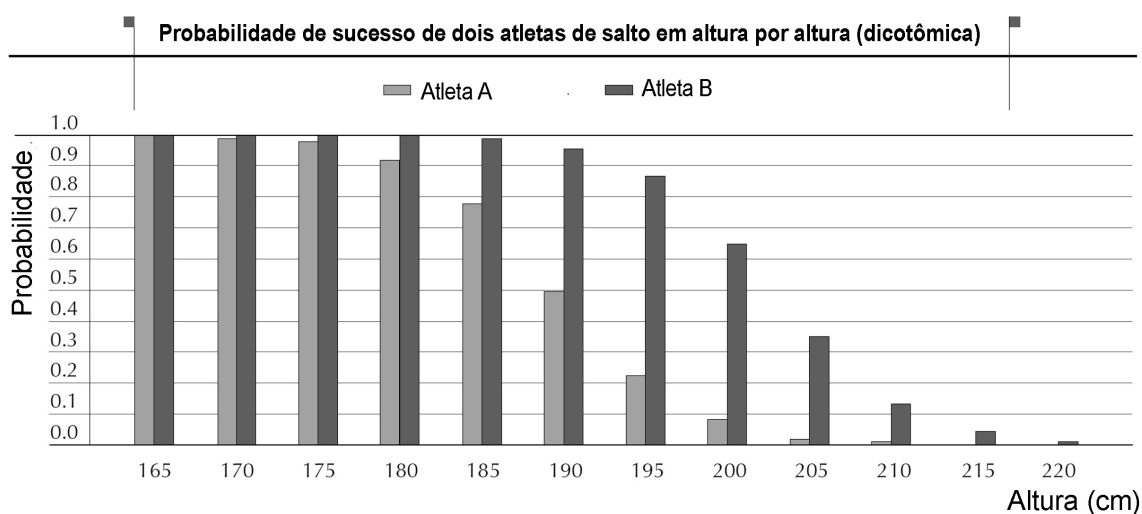


Figura 3.8. Probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano [OECD 2009].

Observe que os dois atletas sempre têm sucesso em 165 cm. A partir daí, a probabilidade de sucesso cai até alcançar 0 para ambos quando a altura a ser saltada é de 225 cm. Para o primeiro atleta, essa diminuição da probabilidade se inicia a altura de 170 cm, enquanto para o segundo, o mesmo fato só ocorre a partir de 185 cm.

Este gráfico apresenta uma imagem visual que faz pensar em transformar a variável dicotômica (sim/não, sucesso/fracasso, 1/0) em uma variável contínua. Em outras palavras, se é traçada uma linha imaginária da medida da probabilidade de sucesso em cada altura (no alto de cada barra, como mostrado na Figura 3.9), pode-se observar uma diferença entre as duas curvas. A variável contínua é a altura, não mais pensada como um sarrafo colocado no alto; e ela

indicará o sucesso ou fracasso de um atleta em particular em função da altura do salto. Por exemplo, percebe-se da curva que o atleta A na altura de 1,90 m tem 50% de probabilidade de ter sucesso no salto, enquanto o atleta B, na altura de 1,90 m, tem probabilidade de 95% de ter sucesso. Com isso, pode-se estimar a probabilidade de sucesso em qualquer altura.

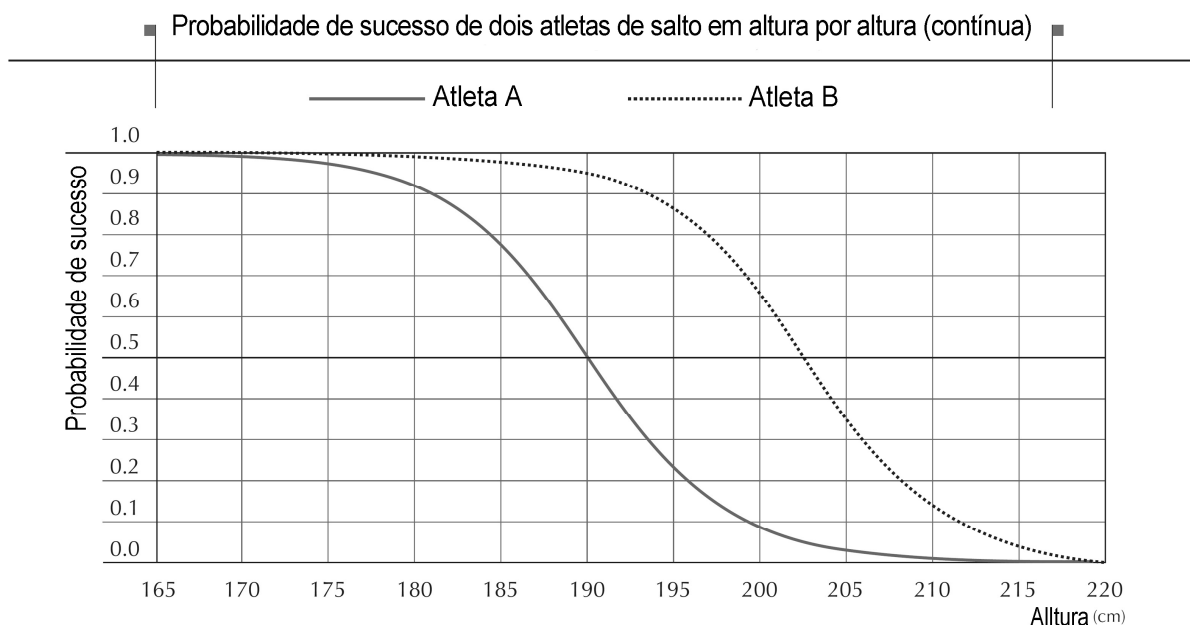


Fig. 3.9. Probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano [OECD 2009].

Estas duas curvas representam a probabilidade de sucesso para os dois atletas. A curva sólida representa a probabilidade de sucesso do atleta A e pontilhada do atleta B.

Pode-se definir o nível de desempenho do atleta como a altura em que a probabilidade de sucesso é 0,5. Isso faz sentido uma vez que abaixo deste nível a probabilidade de sucesso passa a ser menor do que a de fracasso e acima dele, o inverso.

Neste exemplo em particular, o nível de desempenho dos dois atletas é respectivamente 190 cm e 202,5 cm. Observe no diagrama acima que o nível de desempenho do atleta A é visto diretamente no gráfico, enquanto o do atleta B é estimado do modelo. Uma propriedade fundamental deste tipo de abordagem é que o nível (ou seja, a altura) do sarrafo a ser saltado e o desempenho dos atletas são expressas na mesma medida ou escala.

A mesma idéia até aqui explorada está por trás do modelo de Rasch (um tipo de modelo logístico de um parâmetro) para a TRI. A dificuldade dos itens em um teste é análoga à dificuldade do salto com base na altura da barra. Além disso, assim como um salto em particular possui dois resultados possíveis (sucesso ou fracasso), a resposta de um aluno a um determinado item também possui duas possibilidades (acerto ou erro). Da mesma forma que o desempenho do atleta foi definido como a altura para a qual a probabilidade de sucesso no salto é de 0,5, o desempenho ou aptidão do aluno é medida no ponto onde sua probabilidade de sucesso no item é de 0,5.

Uma das características importantes do modelo de Rasch é que ele cria um contínuo no qual tanto o desempenho do estudante quanto a dificuldade do item estarão localizados em uma função probabilística que relaciona essas duas componentes. Alunos de baixo desempenho ou aptidão e itens fáceis estão localizados na parte inferior do contínuo ou escala enquanto alunos de alto desempenho ou aptidão e itens difíceis estarão localizados na parte superior do contínuo ou escala. Na Figura 3.10 está representada a probabilidade de sucesso e de fracasso num item de dificuldade zero.

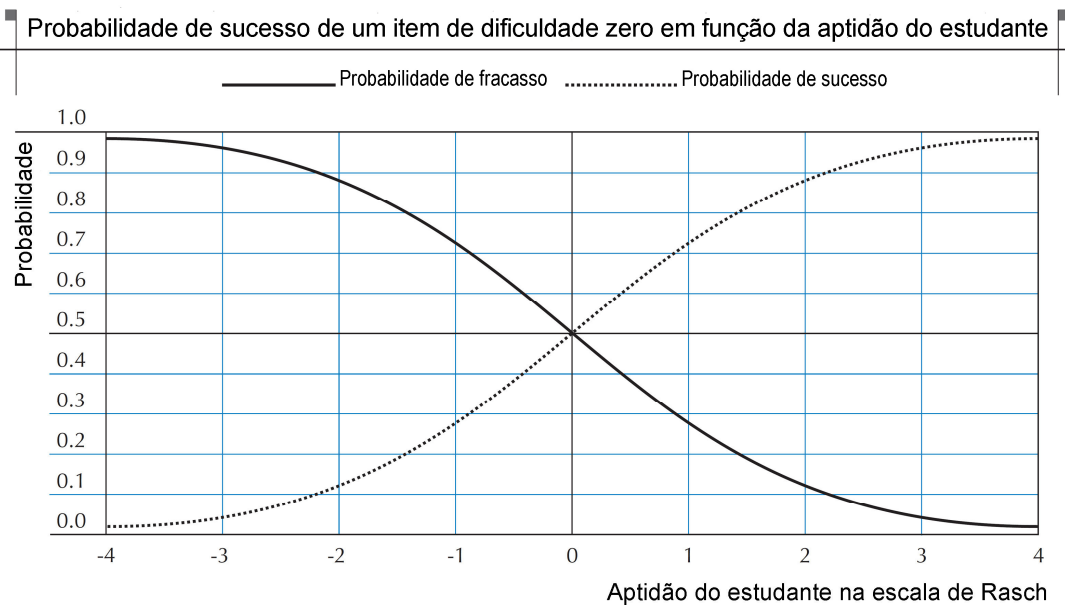


Figura 3.10. probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano [PISA 2009].

Como se pode observar da Figura 3.10, um estudante com uma aptidão zero tem a probabilidade de 0,5 de sucesso ou fracasso num item de dificuldade zero. Um estudante com aptidão -2 tem a probabilidade de um pouco mais de 0,10 de sucesso e um pouco menos de 0,90 de fracasso no mesmo item de dificuldade zero. Mas este mesmo estudante teria probabilidade 0,5 de sucesso num item de dificuldade -2 .

Do ponto de vista matemático, a probabilidade de um estudante com uma aptidão " θ " responder corretamente um item de dificuldade " b " pode ser descrita como:

$$\text{"odds ratio"} = \frac{\text{probabilidade (acerto)}}{\text{probabilidade (erro)}} = \frac{P}{1-P} \quad (\text{equação 3.1})$$

$$\frac{P}{1-P} = e^{(\theta-b)} \quad \rightarrow \quad P(\theta) = \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} = \frac{1}{1+e^{-(\theta-b)}} \quad (\text{equação 3.2})$$

De forma similar, a probabilidade de fracasso $P(\theta) = 0$ é dada por:

$$P(\theta_{\text{fracasso}}) = 1 - \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} \\ \rightarrow P(\theta_{\text{fracasso}}) = \frac{1+e^{(\theta-b)} - e^{(\theta-b)}}{1+e^{(\theta-b)}} = \frac{1}{1+e^{(\theta-b)}} \quad (\text{equação 3.3})$$

Observe que $P(\theta) + P(\theta_{\text{fracasso}}) = 1$, ou seja:

$$P(\theta) + P(\theta_{\text{fracasso}}) = \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} + \frac{1}{1+e^{(\theta-b)}} = \frac{1+e^{(\theta-b)}}{1+e^{(\theta-b)}} = 1$$

A expressão (3.3) para a probabilidade de sucesso é um modelo matemático, o modelo de Rasch, que descreve a probabilidade de sucesso como função da aptidão do aluno – exatamente a ideia de uma curva característica do item para um item de um teste.

Pode-se continuar esta discussão tomando alguns exemplos deste modelo, calculando-se a probabilidade de sucesso para aptidões e níveis de dificuldades estabelecidos.

1º exemplo) Cálculo da probabilidade de sucesso quando a habilidade do estudante é igual a dificuldade do item.

Tabela 3.2. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é igual a dificuldade do item [OECD 2009].

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-2	0,5
-1	-1	0,5
0	0	0,5
1	1	0,5
2	2	0,5

$$\theta - b = 0 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(0)}} = \frac{1}{1+1} = 0,5$$

Logo, a probabilidade de fracasso será: $P(\theta_{\text{fracasso}}) = 1 - 0,5 = 0,5$

2º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é uma unidade menor do que a dificuldade do item.

Tabela 3.3. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é uma unidade menor do que a dificuldade do item [OECD 2009].

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-1	0,27
-1	0	0,27
0	1	0,27
1	2	0,27
2	3	0,27

$$\theta - b = -1 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(-1)}} = \frac{1}{1 + (2,718)^1} = 0,27$$

Logo, a probabilidade de fracasso será: $P(\theta_{\text{fracasso}}) = 1 - 0,27 = 0,73$

3º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é uma unidade maior do que a dificuldade do item.

Tabela 3.4. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é uma unidade maior do que a dificuldade do item [OECD 2009].

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-3	0,73
-1	-2	0,73
0	-1	0,73
1	0	0,73
2	1	0,73

$$\theta - b = 1 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-1}} = \frac{1}{1 + (2,718)^{-1}} = 0,73$$

Logo, a probabilidade de fracasso será: $P(\theta_{\text{fracasso}}) = 1 - 0,73 = 0,27$

4º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é duas unidades menor do que a dificuldade do item.

Tabela 3.5. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é duas unidades menor do que a dificuldade do item [OECD 2009].

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	0	0,12
-1	1	0,12
0	2	0,12
1	3	0,12
2	4	0,12

$$\theta - b = -2 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(-2)}} = \frac{1}{1 + (2,718)^2} = 0,12$$

Logo, a probabilidade de fracasso será: $P(\theta_{\text{fracasso}}) = 1 - 0,12 = 0,88$

5º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é duas unidades maior do que a dificuldade do item.

Tabela 3.6. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é duas unidades maior do que a dificuldade do item [OECD 2009].

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-4	0,12
-1	-3	0,12
0	-2	0,12
1	-1	0,12
2	0	0,12

$$\theta - b = 2 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-2}} = \frac{1}{1 + (2,718)^{-2}} = 0,88$$

Logo, a probabilidade de fracasso será: $P(\theta_{\text{fracasso}}) = 1 - 0,88 = 0,12$

Observe que:

- se a aptidão do estudante é igual ao nível de dificuldade do item, a probabilidade de sucesso será sempre igual a 0,5, independentemente da posição da aptidão e da dificuldade do item no contínuo;
- se a dificuldade do item excede aptidão do estudante em uma unidade, a probabilidade de sucesso será sempre igual a 0,27, independentemente da posição da aptidão e da dificuldade do item no contínuo;

- se aptidão do estudante excede a dificuldade do item em uma unidade, a probabilidade de sucesso será sempre igual a 0,73, independentemente da posição da aptidão e da dificuldade do item no contínuo;
- se duas unidades separam a aptidão do estudante e o nível de dificuldade do item, a probabilidade de sucesso será igual a 0,12 e 0,88, respectivamente, independentemente da posição da aptidão e da dificuldade do item no contínuo.

Desses exemplos, fica evidente que somente a distância entre a aptidão do estudante e o nível de dificuldade do item, no contínuo de Rasch, é fator determinante da probabilidade de sucesso deste aluno na execução de um determinado item.

Os exemplos também ilustram a simetria das escalas. Se um estudante possui aptidão uma unidade abaixo do nível de dificuldade do item, a probabilidade de sucesso deste será de 0,27, ou seja, 0,23 abaixo da probabilidade de sucesso quando essa aptidão e dificuldade do item são iguais. Agora, se ao contrário, o estudante possuir aptidão uma unidade acima do nível de dificuldade do item, sua probabilidade de acerto será de 0,73, ou seja, 0,23 acima da probabilidade de sucesso quando essa aptidão é igual ao nível de dificuldade do item. Da mesma forma, a diferença é de duas unidades gerará um fator de simetria igual a 0,38.

3.3.4. Modelos de TRI

Os modelos existentes para a TRI diferenciam-se basicamente pelo número de parâmetros utilizados para descrever os itens. Os modelos mais populares [Hambleton 1991] são os modelos logísticos de um, dois e três parâmetros, sendo que estes são apropriados para dados de respostas de itens dicotômicos.

3.3.4.1. Modelo Logístico de um Parâmetro

Este modelo para a TRI é um dos mais utilizados. Sua curva característica é dada por

$$P_i(\theta) = \frac{1}{1 + e^{-y(\theta)}} = \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = \theta - b_i \text{ com } i = 1, 2, \dots, n,$$

Nesta expressão,

i refere-se a um item de um total de n itens que constituem o teste; esses itens são dicotômicos, isto é, só admitem respostas do tipo certo (1) ou errado (0); $P_i(\theta)$ é a probabilidade que um examinando, escolhido aleatoriamente, com aptidão θ (parâmetro do modelo), responda ao item i corretamente; a probabilidade tem sempre um valor entre 0 e 1;

b_i é um parâmetro do modelo, que indica o nível de dificuldade do item.

Observa-se que $y(\theta) = \theta - b_i$ corresponde a uma função linear, com coeficiente angular igual a 1. Pode-se verificar que, quando $y(\theta = b_i) = 0$, a probabilidade de acerto $P_i(\theta = b_i) = 0,5$, ou seja, b_i define a aptidão para a qual o candidato tem probabilidade 50% de acertar o item.

Essa parametrização corresponde ao modelo de Rasch, sendo que esse modelo logístico para a TRI é o modelo utilizado pelo PISA.

Considere-se um exemplo da curva característica do item, da questão de nome “Clareza”, uma unidade com dois itens, integrante da prova de Ciências do PISA na edição de 2000.

Esta questão “Clareza” foi tornada pública pelo PISA⁷. Como em todas as unidades do PISA, uma “questão” é constituída de um texto introdutório

⁷ As questões públicas do PISA podem ser consultadas na página do INEP: <http://portal.inep.gov.br/internacional-novo-pisa-itens>. consultado em

seguido de um ou mais itens. A unidade Claridade tem dois itens associados. Na figura 3.11 está apresentado o texto introdutório.

CLARIDADE

Leia as informações abaixo e responda às questões que se seguem.

DURAÇÃO DO DIA EM 22 DE JUNHO DE 1998

Hoje, enquanto o Hemisfério Norte celebra seu dia mais longo, os australianos viverão o seu dia mais curto.

Em Melbourne*, Austrália, o sol nascerá às 7:36 h e se porá às 17:08 h, totalizando nove horas e 32 minutos de claridade.

Compare o dia de hoje com o dia mais longo do ano no Hemisfério Sul, esperado para 22 de dezembro, quando o sol nascerá às 5:55 h e se porá às 20:42 h (horário de verão), totalizando 14 horas e 47 minutos de claridade.

O presidente da Sociedade de Astronomia, Sr. Perry Vlahos, disse que a existência de mudanças nas estações entre os Hemisférios Norte e Sul estava ligada à inclinação de 23 graus da Terra.

*Melbourne é uma cidade no sul da Austrália a uma latitude de cerca de 38 graus ao Sul do equador.

Figura 3.11. Texto introdutório da questão “Claridade” do PISA 2000 [PISA 2000].

O primeiro item desta questão está no formato de múltipla escolha, referindo-se ao conhecimento de fenômenos astronômicos básicos, no caso, especificamente a explicação para a existência de dia e noite (claridade e escuridão). Seu texto está apresentado na figura 3.12.

QUESTÃO 24: CLARIDADE *S129Q01*

Qual é a afirmação que explica por que a claridade e a escuridão ocorrem na Terra?

- A A Terra gira em torno do seu eixo.
- B O Sol gira em torno do seu eixo.
- C O eixo da Terra é inclinado.
- D A Terra gira em torno do Sol.

Figura 3.12. Texto do item 1 da questão “Claridade” do PISA 2000 [PISA 2000].

A construção das ICC’s deste item pode ser feita através da utilização dos dados brutos do PISA (disponíveis em www.pisa.oecd.org). De acordo com

Barroso e Franco [Barroso e Franco 2008], considerando-se o escore bruto dos participantes do exame como sendo a aptidão, pode-se reconstruir as ICC's para diversas sub-amostras dos participantes. Na figura 3.13 [Barroso e Franco 2008] apresentamos as ICC's para a item "Clareza – Q1" da figura 3.15 para todos os participantes do PISA, do Brasil, do Japão e do Reino Unido.

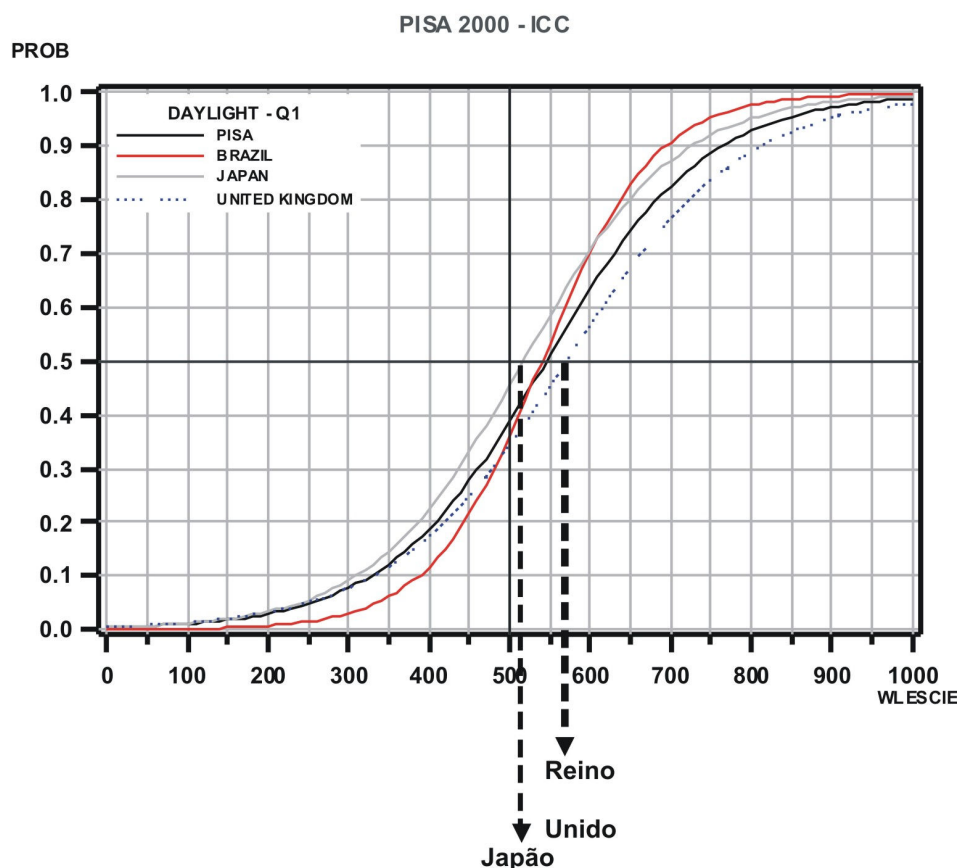


Figura 3.13. ICC de Todos participantes do PISA 2000, do Brasil, do Japão e do Reino Unido [Barroso e Franco 2008].

Observa-se, pela posição em relação ao eixo de aptidão (WLESCIE no gráfico) que esse item foi um pouco mais difícil para os estudantes do Reino Unido (linha tracejada azul) do que para os do Japão (linha contínua cinza), já que a aptidão que corresponde a uma probabilidade de acerto do item igual a 50% foi maior para o Reino Unido do que para o Japão. O índice de dificuldade para os participantes brasileiros é intermediário entre esses dois países e praticamente igual ao da amostra global.

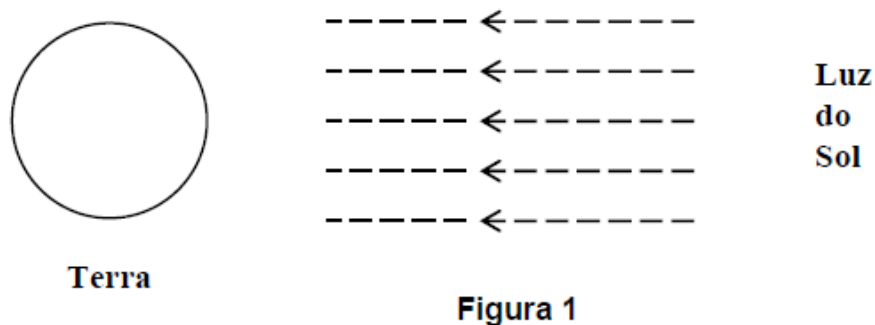
Esta constatação revela a idéia básica da construção das ICC's. O desempenho dos participantes brasileiros (no PISA 2000) revela-se bastante pior

do que do restante do mundo [Barroso e Franco 2008] quando observamos os histogramas dos escores no exame. O índice de acerto dos estudantes brasileiros neste item é muito mais baixo do que do total de participantes e dos participantes do Japão e Reino Unido, por exemplo [Barroso e Borgo 2010]. Mesmo assim, as ICC's são praticamente idênticas.

Na Figura 3.14 mostramos o item 2 da unidade “Clareza”. Este item exige o desenho das posições do eixo da Terra e do Equador, sendo considerada bastante difícil.

QUESTÃO 25: CLARIDADE S129Q02- 0 1 2 8 9

A Figura 1 demonstra os raios de luz do sol se refletindo sobre a Terra.



Suponha que seja o dia mais curto em Melbourne. Mostre o eixo da Terra, o Hemisfério Norte e o Hemisfério Sul na Figura 1.

Figura 3.14. Texto do item 1 da questão “Clareza” do PISA 2000 [OECD 2000].

A construção das ICC's deste item para todos os participantes do PISA e para os participantes do Brasil, Japão, Portugal e Reino Unido [Barroso e Franco 2008] revela que este item apresenta um grande nível de dificuldade. Como se pode observar na Figura 3.15, o índice de dificuldade desse item é de 650, ou seja, apenas os participantes cujo escore foi acima da média em 1,5 desvio padrão (menos de $\cong 10\%$ dos participantes) têm probabilidade de acertar este item superior a 50%.

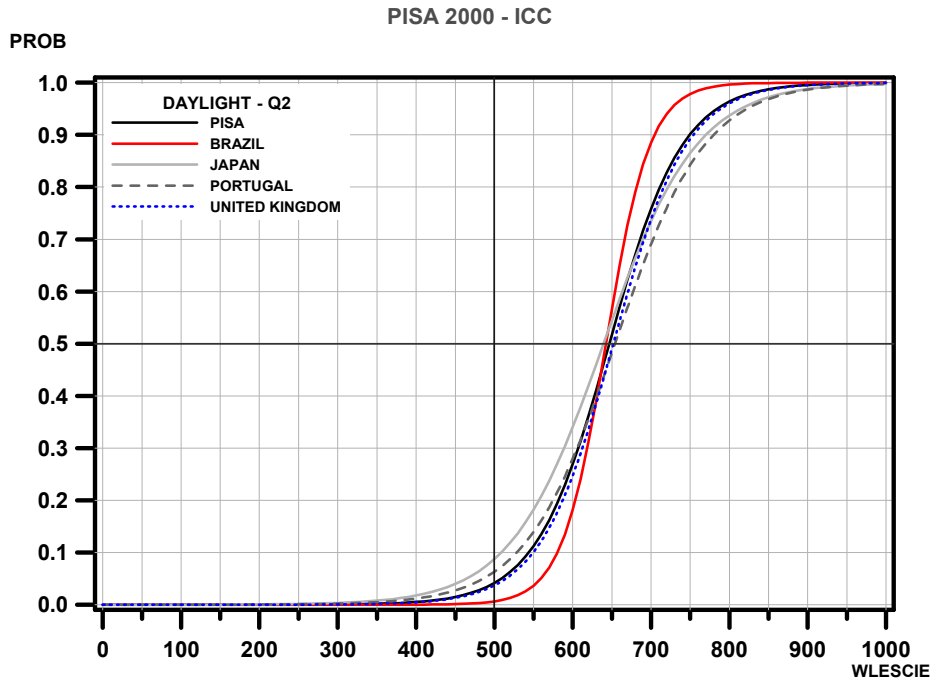


Figura 3.15. ICC de Todos participantes do PISA 2000, do Brasil, do Japão, de Portugal e do Reino Unido [Barroso e Franco 2008].

Ainda se pode observar na figura o índice de dificuldade deste item 2 é praticamente o mesmo para todos os países representados no diagrama.

3.3.4.2. Modelo logístico de dois parâmetros

O modelo logístico de dois parâmetros permite avaliar dois parâmetros do item: a dificuldade e a discriminação. A curva característica deste modelo para TRI é dada pela equação

$$P_i(\theta) = \frac{1}{1 + e^{-y(\theta)}} = \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = Da_i(\theta - b_i) \text{ com } i = 1, 2, \dots, n,$$

onde

D é um fator de escala introduzido em virtude de considerações estatísticas, e vale 1,7 [Pasquali 2009];

a_i representa o parâmetro de discriminação do item. Ele é proporcional à inclinação da ICC no ponto b_i na escala de aptidão e é determinado por

$$\left. \frac{dP_i}{d\theta} \right|_{\theta=b_i} = \frac{D}{4} a_i$$

Esse valor pode variar de 0 a ∞ , mas, normalmente, varia entre 0 e 2. Valores negativos não são aceitos pois representariam que quanto menor a aptidão do estudante, maior a possibilidade dele acertar o item.

A figura 3.16 mostra os parâmetros de dificuldade (b_i) e de discriminação (a_i) de dois itens.

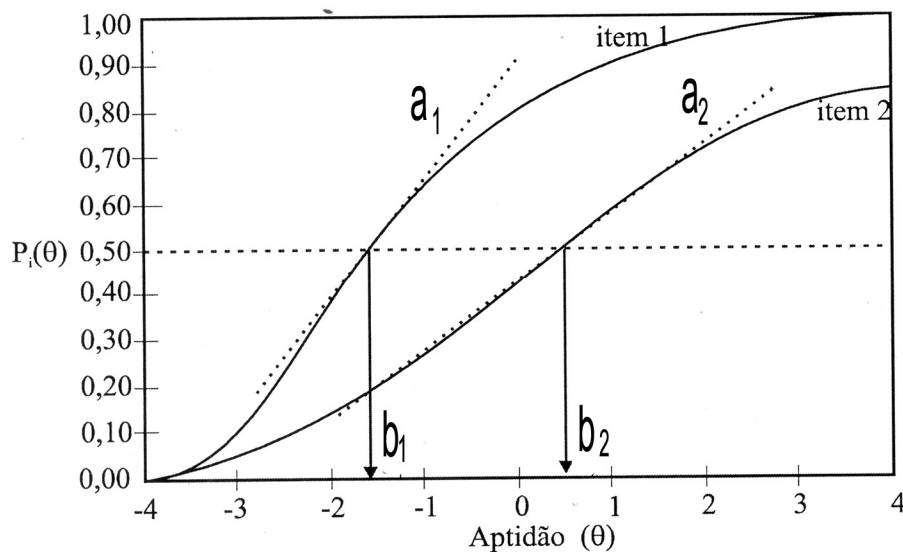


Figura 3.16. Parâmetros de dificuldade (b) e discriminação (a) de dois itens [Pasquali 2009]

Observe que o item 2 apresenta uma maior nível de dificuldade que o item 1 (valor de b maior). No entanto, o item 1 é mais discriminatório, pois possui maior inclinação no ponto correspondente a probabilidade de acerto igual a 0,5.

3.3.4.3. Modelo logístico de três parâmetros

O modelo logístico de três parâmetros permite avaliar dois parâmetros do item: a dificuldade, a discriminação e a resposta dada ao acaso (o “chute”). A curva característica deste modelo para TRI é dada pela equação

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = Da_i(\theta - b_i) \text{ com } i = 1, 2, \dots, n,$$

onde “ c_i ” é o parâmetro do item que permite avaliar a resposta correta dada ao item por acaso e é expresso pela assíntota inferior da curva (Pasquali, 2009), no nível mais baixo do contínuo de aptidões.

O diagrama abaixo apresenta os parâmetros de dificuldade (b_i), de discriminação (a_i) e de “chute” (c_i) de três itens.

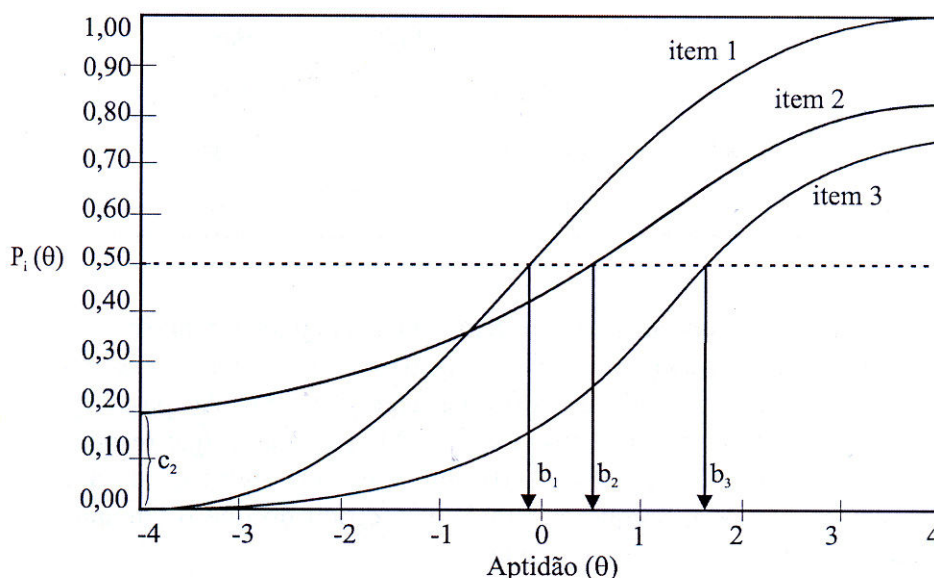


Fig. 3.17. Parâmetros de dificuldade (b) e discriminação (a) de dois itens [Pasquali 2009]

Observe que existe um nível de dificuldade crescente, do item 1 para o 3, que pode ser observado pela posição das ICC (quando $P_i(\theta) = 0,5$) em relação ao eixo das aptidões ($b_1 < b_2 < b_3$). É também possível determinar o índice de discriminação de cada uma das curvas, pela inclinação da mesma nesse mesmo ponto (quando $P_i(\theta) = 0,5$).

O que distingue o modelo de três parâmetros é que no item 2, por exemplo, o ponto da ordenada que é cortado pela ICC indica a existência de “chute”, com probabilidade nesse item de 20%. Observe que nos itens 1 e 2 essa probabilidade é zero. Segundo Pasquali [Pasquali 2009]

“A lógica que fundamenta essa interpretação da assíntota é a seguinte: supostamente o sujeito não tem habilidade praticamente nenhuma, pois ele tem um θ menor que -3 , e apesar disso acerta o item; conseqüentemente, ele só pode ter chutado e teve sorte, porque acertou.” (p. 89)

|

Sendo esse modelo de três parâmetros o modelo adotado para análise das provas do Exame Nacional do Ensino Médio, cabe citar um exemplo dado para o cálculo da probabilidade de acerto de aluno em um item, [Pasquali 2009].

Nos exemplos da tabela, pode-se ter uma ideia de como esse modelo estima a probabilidade de acertar um item, variando os parâmetros dos itens e da aptidão.

Tabela 3.7. Relação dos parâmetros do item e probabilidade de acerto [Pasquali 2009].

Item	Parâmetros				P _i (θ)
	a	b	c	θ	
1	1,00	-3,00	0,00	1,00	1,00
2	1,50	-2,10	0,10	-2,10	0,55
3	2,00	-0,50	0,15	-1,30	0,20
4	2,50	0,00	0,20	0,50	0,91
5	2,10	1,30	0,11	1,00	0,34
6	1,80	2,10	0,25	2,05	0,60
7	1,50	3,00	0,16	2,50	0,34

Observe como foi calculada a probabilidade de acerto do item 6 da tabela 3.7 pelo modelo logístico de 3-parâmetros:

$$P(\theta) = c_6 + (1 - c_6) \frac{e^{D \cdot a_6 (\theta - b_6)}}{1 + e^{D \cdot a_6 (\theta - b_6)}}$$

$$P(\theta) = 0,25 + (1 - 0,25) \frac{2,7183^{1,7 \times 1,8(2,05 - 2,1)}}{1 + 2,7183^{1,7 \times 1,8(2,05 - 2,1)}}$$

$$P(\theta) = 0,60$$

Após essa discussão, pode-se observar que a Teoria de Resposta ao Item se constitui numa poderosa ferramenta que, devido à independência do teste em relação ao examinando, permite a avaliação e a comparação de diferentes grupos com diferentes aptidões. Isso é de fundamental importância para que se possa obter resultados confiáveis que permitam a tomada de ações sejam pelos professores ou elaboradores de políticas públicas.

CAPÍTULO 4 – QUESTIONÁRIO SOBRE A APRENDIZAGEM DE FÍSICA TÉRMICA

Neste capítulo é feita uma discussão sobre a pesquisa em concepções de física térmica, a partir do trabalho desenvolvido por Yeo e Zadnik [Yeo e Zadnik 2001], que utilizaram de um questionário para determinação destas concepções. Após a construção de uma versão em português deste questionário, ele foi aplicado a estudantes no estado do Rio de Janeiro e os resultados obtidos analisados através de estatística tradicional e construção de ICC's.

4.1 – Testes Conceituais para o Ensino de Física

A pesquisa em ensino de física, durante certo período, focalizou suas atividades na identificação dos conceitos não científicos presentes na estrutura cognitiva tanto dos estudantes da educação básica quanto do ensino superior [Arons 1997]. Mais do que isso, constatou-se que esses conceitos não científicos são muito resistentes à mudança, e que são bastante coincidentes em diferentes grupos, países e idades. No tópico de física térmica, o número de trabalhos que indicam as dificuldades na aprendizagem dos conceitos básicos sobre o tema é muito grande [Viennot 1998; Arons 1997], sendo que R. Duit e colaboradores [Duit 2009], durante o período dos anos 70 a 2009, mantiveram um repositório sobre os trabalhos com levantamentos das concepções dos professores e estudantes, permitindo assim a busca desses conceitos.

Alguns desses conceitos não científicos constantemente encontrados pelos pesquisadores entre os estudantes nos cursos introdutórios de física térmica estão relacionados aos conceitos de *calor* (calor é uma substância, não é uma forma de energia; calor e frio são diferentes; calor e temperatura são a mesma coisa; calor não é algo mensurável), de *temperatura* (temperatura é uma forma de calor; a temperatura pode ser determinada pelo toque com a pele; as percepções de frio e calor não estão relacionadas á transferência de calor; um corpo frio não pode conter calor; a maior temperatura que uma substância pode atingir é a sua temperatura de ebulição), das *mudanças de temperatura* e *transferência de calor* (por exemplo, fornecer calor sempre implica em aumentar a temperatura; o calor

sobe; a temperatura pode ser transferida; objetos que estão em contato não atingem necessariamente uma temperatura de equilíbrio), e às *propriedades térmicas dos materiais* (a temperatura é uma propriedade do material ou do objeto; o ponto de ebulição da água é sempre 100° C; vapor de água está sempre a mais de 100° C, entre outras).

O conhecimento a respeito da existência desses conceitos implicou num extenso desenvolvimento de pesquisas que procuravam entender quais as concepções dos estudantes e a elaboração de propostas de intervenção na forma de produção de materiais, sequências e estratégias didáticas para melhorar a aprendizagem.

Para avaliar a efetividade das iniciativas de ensino, algumas propostas de avaliações padrão começaram a ser sugeridas. O nome “inventário de conceitos” (concept inventory) começou a ser utilizado para referir-se a testes padronizados, preparados especificamente para avaliar se um estudante teria um conhecimento operacional preciso de um conjunto específico de conceitos. O primeiro teste produzido na área de ensino de Física, o “Force Concept Inventory” (FCI), elaborado por Hestenes, Halloun e Wells [Hestenes 1992], foi amplamente disseminado e é utilizado até os dias de hoje. Outros testes foram propostos e vem sendo utilizados, em mecânica, astronomia, eletricidade, magnetismo, ótica e muitos outros.

4.2- Um Teste para Avaliar Conceitos Introdutórios de Física Térmica

No tópico de Física Térmica, Yeo e Zadnik [Yeo e Zadnik 2001] propuseram um questionário que permitisse a avaliação da evolução dos conceitos entre alunos de 15 a 19 anos. A partir de um levantamento sobre as concepções dos estudantes, foi construído um questionário com 26 itens objetivos. A versão em português aplicada aos estudantes está apresentada no Apêndice 1. No Apêndice 2, apresenta-se a relação elaborada da relação entre os itens e as respostas (as corretas e os distratores) com as concepções levantadas.

Na apresentação do questionário, os autores [Yeo e Zadnik 2001] afirmam que as respostas dadas pelos alunos correspondem às suas concepções, pois os itens do questionário são apresentados em contextos comuns do cotidiano e de sala de aula, dificultando respostas correspondentes àquelas ensinadas em sala de aula, às quais muitas vezes o aluno repete sem acreditar.

Nesse contexto, os autores afirmam que há duas maneiras pelas quais as concepções alternativas podem afetar a escolha da resposta do item pelo aluno [Yeo e Zadnik 2001]:

“A primeira, quando a crença na concepção alternativa é tão forte que induz o estudante ao enxergar a alternativa “correta” como não plausível. Por exemplo, um aluno que acredita que o plástico é “naturalmente mais quente que o metal” não pode aceitar que um corpo de plástico e um de metal, retirados dentro de um refrigerador, estejam à mesma temperatura (ver questão 9). A segunda, quando a concepção alternativa faz com que a alternativa incorreta pareça a mais plausível de todas. Por exemplo, um aluno que acredita que a temperatura de ebulição da água não possa permanecer constante pode selecionar como 110°C a temperatura da água que está fervendo por algum tempo (ver questão 5).”¹ (p.499)

Uma versão em português da tabela onde esses autores estabeleceram a possível correspondência entre cada item e a sua concepção alternativa é apresentada na Tabela 4.1.

Tabela 4.1. Relação entre as concepções não-científicas e itens do questionário.

Concepções dos Estudantes	nº dos itens
A. Concepções sobre calor	
1. Calor é uma substância.	10, 22
2. Calor não é energia.	22
3. Calor e frio são diferentes, em vez de extremos opostos de um contínuo.	10, 13, 18, 23, 24
4. Calor e temperatura são a mesma coisa.	15, 18
5. Calor é proporcional a temperatura.	7, 11, 15
6. Calor não é um conceito mensurável ou quantificável.	7

¹ Original: “First, a firmly held alternative conception may prompt students to dismiss the “correct” response as implausible. For example, a student who believes that plastic is “naturally warmer than metal” will not accept that plastic and metal containers taken from the refrigerator can be at the same temperature (see question 9). Second, a different alternative conception may make an incorrect response seem the most plausible one. For example, a student who believes that the temperature of boiling water cannot remain constant might select 110°C as the temperature of water that has been boiling for some time (see question 5).”

<p>B. Concepções sobre calor</p> <p>7. Calor é uma substância.</p> <p>8. Calor não é energia.</p> <p>9. Calor e frio são diferentes, em vez de extremos opostos de um contínuo.</p> <p>10. Calor e temperatura são a mesma coisa.</p> <p>11. Calor é proporcional a temperatura.</p> <p>12. Calor não é um conceito mensurável ou quantificável.</p>	<p>10, 22</p> <p>22</p> <p>10, 13, 18, 23, 24</p> <p>15, 18</p> <p>7, 11, 15</p> <p>7</p>
<p>C. Concepções sobre temperatura</p> <p>13. A temperatura é a “intensidade”de calor (<i>ou medida da quantidade de calor</i>).</p> <p>14. A temperatura pode ser determinada pelo toque ou pele</p> <p>15. As percepções de calor e temperatura não estão relacionados a transferência de energia.</p> <p>16. Quando a temperatura de ebulição permanece constante, há algo errado.</p> <p>17. A temperatura de ebulição é a maior temperatura que um corpo pode atingir.</p> <p>18. Um corpo frio não contém calor.</p> <p>19. A temperatura de um objeto depende do seu tamanho.</p> <p>20. Não há limite para a temperatura mais baixa.</p>	<p>15</p> <p>16</p> <p>10, 18, 21, 22</p> <p>5</p> <p>19</p> <p>7, 10, 11, 22, 26</p> <p>1, 9, 14</p> <p>25</p>
<p>D. Concepções sobre transferência de calor e mudança de temperatura</p> <p>21. Aquecimento sempre resulta em um aumento de temperatura.</p> <p>22. O calor só se move para cima</p> <p>23. Calor sobe.</p> <p>24. Calor e frio fluem como líquidos.</p> <p>25. Temperatura pode ser transferida.</p> <p>26. Objetos com diferentes temperaturas em contato entre si, ou em contato com o ar em diferente temperatura, não necessariamente atingem a mesma temperatura (não possuem o conceito de equilíbrio térmico)</p> <p>27. Objetos quentes esfriam naturalmente e objetos frios esquentam naturalmente.</p> <p>28. Calor flui mais lentamente através de condutores fazendo-os parecer mais quentes.</p> <p>29. A teoria cinética não explica de fato a transferência de calor (os conceitos são expressos mas não se acredita neles).</p>	<p>3, 4, 5</p> <p>20</p> <p>20</p> <p>10, 13</p> <p>7, 13</p> <p>1, 2, 3, 6, 9, 10, 17, 24</p> <p>3, 13</p> <p>25</p> <p>18, 20, 21</p>

Yeo e Zadnik testaram, validaram e posteriormente aplicaram esse questionário a 483 estudantes australianos do 10º ao 13º grau. Na prática, tem-se que o 10º grau corresponde aos alunos (de 15 a 16 anos) em que o conhecimento dos conceitos de física térmica provém de um currículo geral de formação em ciências. No 11º grau os estudantes (de 16 a 17 anos) já foram expostos aos

conceitos de calor, temperatura, transferência de calor, calor específico e calor latente. No 12º grau (17 a 19 anos), os alunos, além de aulas de física, assistem aulas de química, onde também são abordados conceitos de física térmica. Finalmente, no 13º grau, que corresponde ao primeiro ano da universidade (além dos estudantes já terem passado pelo 11º e 12º graus) estudam novamente conceitos de física térmica. Os resultados deste trabalho podem ser encontrados no artigo de Yeo e Zadnik [Yeo e Zadnik 2001]².

4.3- Aplicação do Teste a Estudantes do Ensino Médio do Rio de Janeiro

A experiência de anos de trabalho com alunos de ensino médio e troca de experiências com outros professores evidenciam a reincidência de observações e relatos sobre o ensino dos conceitos de física térmica, apontando para o fato de que os alunos quase sempre apresentavam os mesmos tipos de dúvidas e os mesmos erros conceituais nas avaliações que envolviam esse tema.

Na busca de se melhorar a aprendizagem desses conceitos, pode-se utilizar, como estratégia, a construção de materiais didáticos e a elaboração de ações para tentar superar as dificuldades de compreensão. No entanto, é necessário acompanhar a aplicação e o desenvolvimento dessas ferramentas, sendo assim útil a criação de mecanismos que sejam capazes de avaliá-las. Com este intuito, aplicou-se a alunos brasileiros a versão em português, apresentada no Apêndice 1, do questionário proposto por Yeo e Zadnik.

Esse questionário foi aplicado durante o ano de 2009 a 433 alunos do ensino médio regular, turno diurno, de duas escolas da rede privada no Estado do Rio de Janeiro. As famílias desses estudantes apresentam um bom poder aquisitivo, permitindo que estes estudassem sem a necessidade de exercer atividades remuneradas.

As informações relativas à aplicação do questionário por escola e por série são apresentadas nas Tabelas 4.2 e 4.3.

² Disponível também em:
<http://staff.science.uva.nl/~eberg/Antwerpen/Diagnostische%20toetsen/Heat%20and%20TemperatureTPT%20YeoZadnik.pdf>. Acesso em 25/06/2009.

Tabela 4.2. Distribuição dos questionários por escola.

	Número de respondentes	Percentual
Colégio 1	273	63
Colégio 2	160	37
Total	433	100

Tabela 4.3. Distribuição dos questionários por série.

	Número de respondentes	Percentual
1ª série	178	41,1
2ª série	129	29,8
3ª série	126	29,1
Total	433	100

Além disso, foi feita uma separação dos alunos de acordo com o nível de instrução formal dos conceitos de física térmica. Para isso, foram divididos em grupos denominados de “pré-instrução”, quando não haviam estudado o tema de física térmica na educação básica, “pós-instrução”, quando já haviam estudado o tema e finalmente de “pós-instrução com repetição” para aqueles que já haviam feito a revisão destes conceitos. A distribuição dos questionários aplicados em cada uma dessas categorias está apresentada na Tabela 4.4.

Tabela 4.4. Nível de instrução formal dos conceitos de física térmica.

	Número de respondentes	Percentual
pré-instrução	111	25,6
pós-instrução	196	45,3
pós-instrução com repetição	126	29,1
Total	433	100

4.4- Metodologia de Análise das Respostas

A análise dos questionários aplicados foi feita com base na estatística tradicional (percentual de acertos por questão e obtenção dos escores brutos) seguida da construção das curvas características dos itens [Hambleton 1991; Hambleton 1993]. Estas curvas características constituem a base da Teoria da Resposta ao Item, discutida no Capítulo 3.

Como anteriormente citado, na TRI, para cada item, prevê-se que a probabilidade de acerto é dada por uma função $P_i(\theta)$ da aptidão θ do respondente. Um dos modelos possíveis para esta função é (no modelo de dois parâmetros) que é determinado pela relação:

$$P_i(\theta) = \frac{1}{1 + e^{-y(\theta)}}, \text{ com } y(\theta) = Da_i(\theta - b_i) \quad i = 1, 2, 3, \dots, n. \quad (\text{equação 4.1})$$

Nesta expressão, θ representa a aptidão do aluno, b_i é um parâmetro que determina a dificuldade do item e a_i é o parâmetro de discriminação do item. Estes parâmetros, de acordo com a TRI, correspondem: no caso da dificuldade do item, ao valor de θ para o qual a probabilidade de acerto do item é de 50%, ou seja, igual a 0,5 e no caso da discriminação do item, à inclinação da curva no ponto correspondente a probabilidade de acerto igual a 50%.

Não foi feito nenhum cálculo de TRI neste trabalho; apenas utilizou-se a ideia de construção empírica das Curva Característica dos Itens para as diversas questões apresentadas no questionário. Utilizou-se como aptidão do aluno a nota (escore bruto, a soma dos acertos no teste) para a construção e avaliação das curvas características para cada um dos itens do teste.

4.5- Apresentação e Análise dos Resultados

Inicialmente, fez-se uma comparação entre os resultados dos questionários obtidos com estudantes australianos e brasileiros, e construiu-se o gráfico da média dos resultados dos alunos para cada um dos diferentes grupos, levando-se em consideração uma pontuação máxima possível, que corresponde ao número total de itens do questionário, 26 pontos. Este resultado está apresentado na Figura 4.1. Na Figura 4.2, reproduz-se o resultado apresentado para alunos australianos [Yeo e Zadnik 2001].

Pode-se observar dessas figuras que o resultado australiano e o obtido a partir dos questionários aplicados aos alunos brasileiros apresentam semelhanças que apontam para uma melhora das notas obtidas à medida que os estudantes entram em contato com o estudo formal dos conceitos de física térmica.

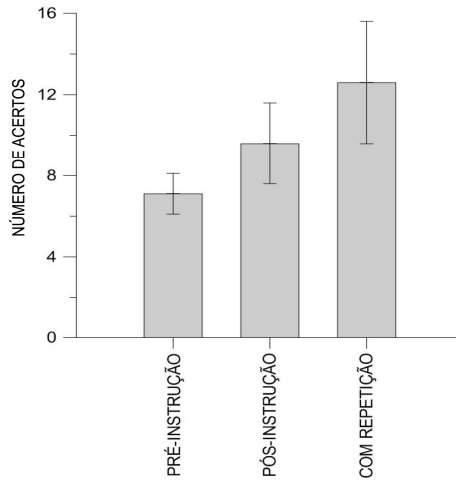


Figura 4.1. Média de acertos obtidos pelos estudantes no Brasil

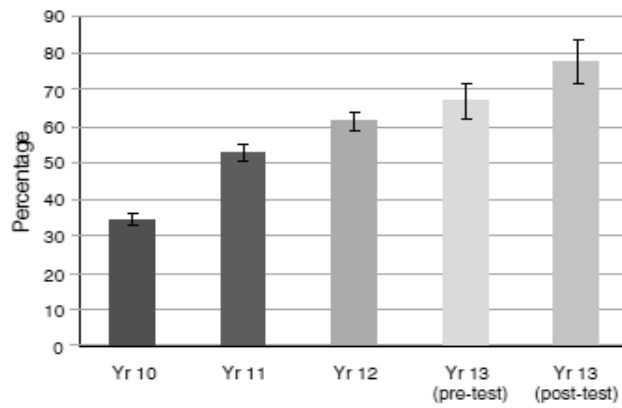
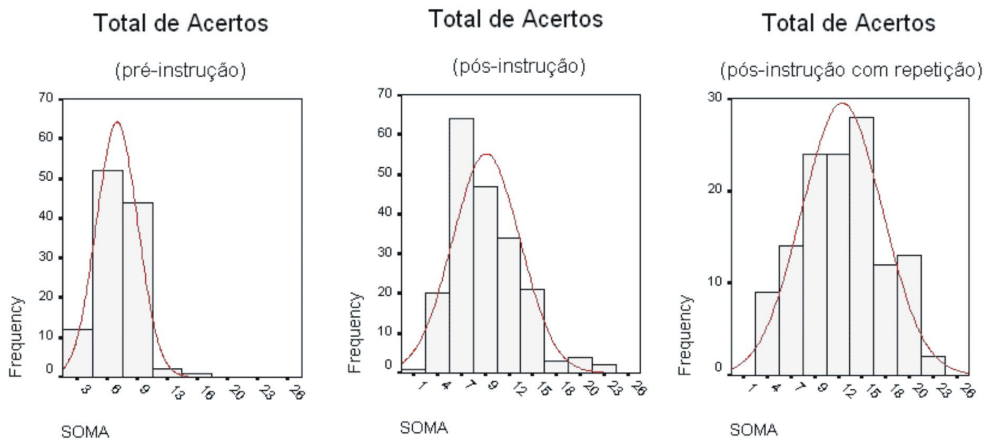


Figura 4.2. Média de acertos obtidos pelos estudantes da Austrália [Yeo e Zadnik 2001].

Na Figura 4.3, apresentam-se os histogramas dos resultados (total de acertos) dos respondentes no Brasil de acordo com a instrução formal recebida pelos estudantes, onde a melhora do desempenho destes estudantes à medida que eles recebem a instrução formal é ratificada.



Fig

ura 4.3. Total de acertos de acordo com o nível de instrução

Buscando-se ter uma percepção do nível de dificuldade do questionário, na Figura 4.4 apresenta-se o histograma do total de acertos para todos os alunos.

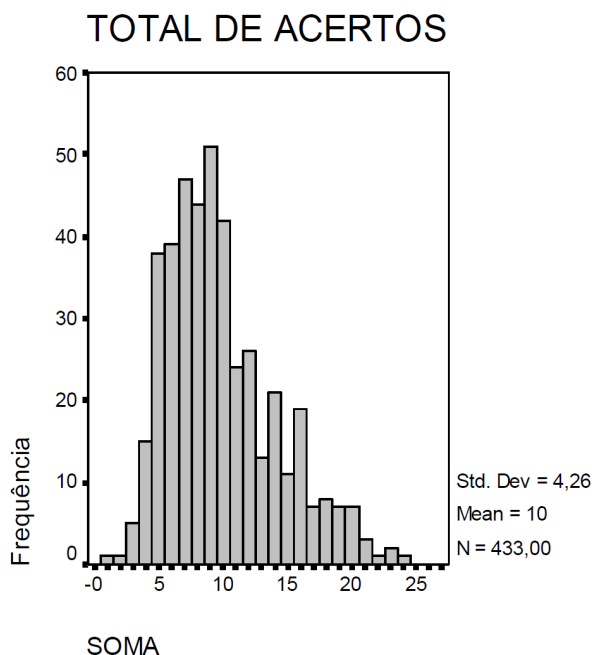


Figura 4.4. Resultado (escore bruto) da aplicação do questionário

Observa-se que a média de acertos para todos os estudantes é 10 acertos (nota 3,6 em 10) com desvio padrão 4,3 (1,6 em 10). Esses valores obtidos indicam que o questionário, como um todo, apresentou um nível de dificuldade de médio para alto uma vez que esta média representa um total de acertos inferior a 50% do escore total do questionário³.

Apesar da existência de uma discreta melhora no desempenho dos alunos, a partir de aulas no formato tradicional (aulas expositivas, exercícios e trabalhos), observa-se que o questionário apresenta este nível de dificuldade mesmo quando os resultados são separados por níveis de instrução, como se pode observar na Tabela 4.5.

Tabela 4.5. Média de acertos das questões por nível de instrução.

	Acertos (em 26)	Nota (em 10)
Pré-Instrução	7 ± 2	2,7 ± 0,9
Pós-Instrução	10 ± 4	3,8 ± 1,5
Pós-Instrução com repetição	13 ± 5	5,0 ± 1,8

³ Foram construídos os gráficos dos colégios individualmente; o resultado se mantém para cada um deles, mesmo com o perfil dos dois colégios sendo diferentes.

Fazendo-se o ajuste das respostas dos alunos por regressão logística (modelo de 2 parâmetros)⁴, tomando-se a aptidão como sendo o escore bruto dos estudantes no teste, foi possível determinar o índice de dificuldade (b_i) de cada questão e o indicador de discriminação de cada item, como apresentado na Tabela 4.6.

Tabela 4.6. Índice de dificuldade e discriminação por questão

Questão	Índice de dificuldade	Índice de discriminação	Questão	Índice de dificuldade	Índice de discriminação
Q01	7,92	0,031	Q14	14,26	0,115
Q02	11,64	0,054	Q15	21,35	0,026
Q03	9,49	0,085	Q16	13,34	0,069
Q04	22,61	0,017	Q17	18,13	0,043
Q05	13,70	0,043	Q18	15,50	0,035
Q06	23,34	0,025	Q19	13,33	0,053
Q07	3,50	0,020	Q20	15,36	0,045
Q08	16,81	0,050	Q21	16,03	0,029
Q09	16,38	0,045	Q22	9,32	0,05
Q10	7,88	0,077	Q23	5,71	0,079
Q11	15,64	0,031	Q24	15,24	0,069
Q12	11,02	0,031	Q25	13,40	0,070
Q13	9,15	0,105	Q26	10,87	0,051

Nas Figuras 4.5 e 4.6, apresenta-se graficamente o valor do índice de dificuldade e de discriminação de cada item para esses alunos. A partir desses gráficos, pode-se identificar as questões que foram mais difíceis para os respondentes e identificar as que possuem maior capacidade de discriminação.

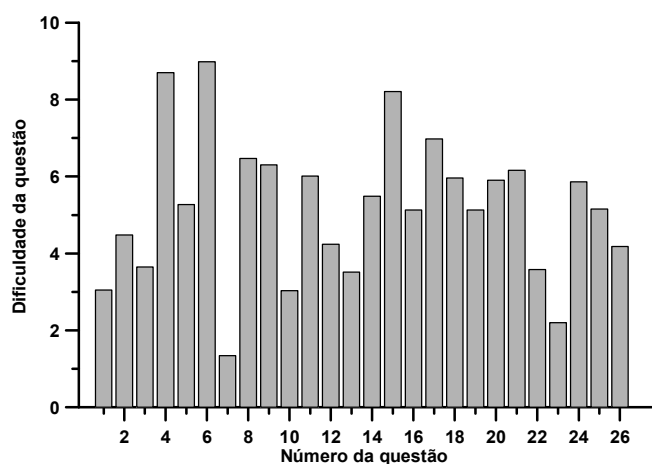


Figura 4.5. Indicador de dificuldade (b_i) por questão.

⁴ Utilizou-se a regressão logística do SPSS.

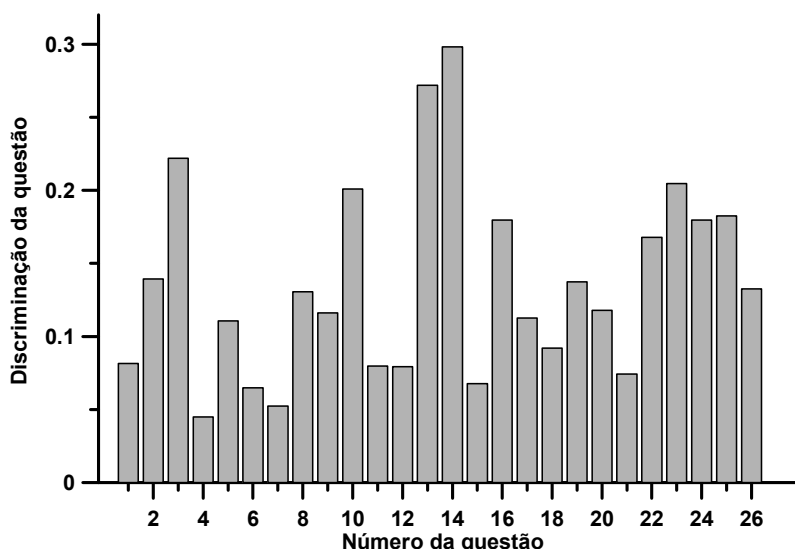


Figura 4.6. Indicador do nível de discriminação da questão.

A seguir, são apresentados os textos das questões 4 e 6, suas tabelas de freqüências e curvas características. Como se observa na Figura 4.5, estes são os dois itens que apresentam o maior índice de dificuldade para os respondentes.

Questão 4: Sobre o fogão está uma chaleira cheia de água. A água começou a ferver rapidamente. A temperatura mais provável da água é: (a) 88°C; (b) 98 °C; (c) 110 °C; (d) Nenhuma das respostas acima pode estar correta.

Questão 6: Qual você acha que é a temperatura do vapor formado, acima da água fervente, na chaleira? (a) 88°C; (b) 98 °C; (c) 110 °C; (d) 120 °C.

Na Tabela 4.7, apresenta-se a estatística descritiva das respostas a esta questão discriminada por nível de instrução.

Tabela 4.7. Tabela de freqüência das respostas às alternativas da questão 4.

alternativas		pré	pós	rep	total
a	88°C	25,23	12,24	15,87	16,63
b	98°C	26,13	29,59	32,54	29,56
c	110°C	26,13	32,65	24,60	28,64
d	Nenhuma das respostas acima podem estar corretas.	22,52	24,49	26,98	24,71
8	Resposta inválida	0,00	1,02	0,00	0,46
9	Reposta faltando	0,00	0,00	0,00	0,00
		100,00	100,00	100,00	100,00

Os dados obtidos permitem a construção da curva da probabilidade de acerto deste item empiricamente, e a curva característica do item a partir da

equação do modelo de 2 parâmetros. Esta curva está apresentada na Figura 4.7. O mesmo é feito para a Questão 6, com a Tabela 4.8 e a Figura 4.8.

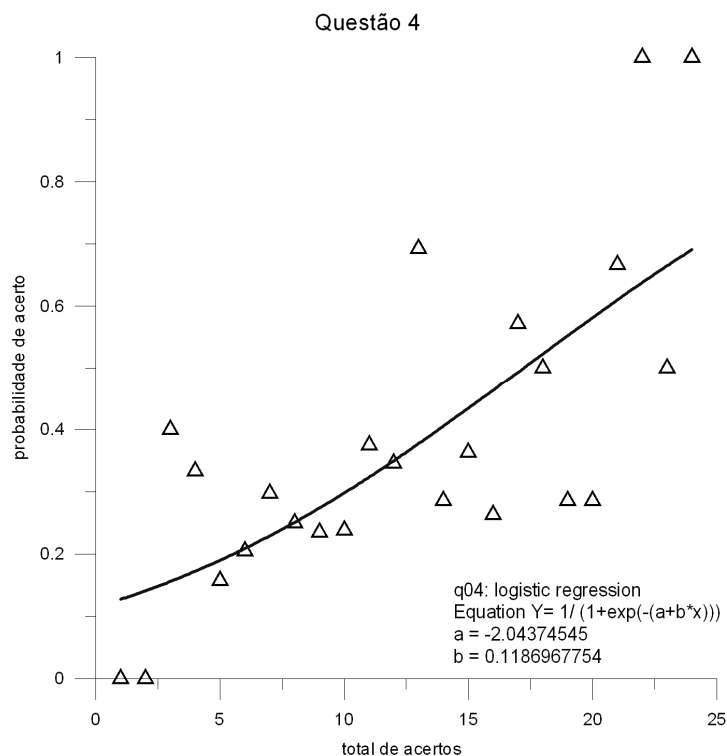


Figura 4.7. Curva Característica do Item da questão 4.

Na questão 4, 53% dos alunos marcaram os distratores “c” ou “d”. Os 25% que marcaram a alternativa “d”, que afirma não existir resposta correta, corroboram a afirmação que os alunos acreditam na idéia de que a água entra em ebulição obrigatoriamente a 100°C. Os 29% que marcaram a alternativa “c”, que afirma que o mesmo fenômeno ocorre a 110°C, juntamente com os 57% dos respondentes da questão 6 que marcaram para a temperatura do vapor da água um valor superior a 100°C, evidenciando que o vapor acima da água fervente deve possuir uma temperatura superior a esta, evidenciam, nos estudantes, a não compreensão da lei zero da termodinâmica, ou seja, o não entendimento do conceito de equilíbrio térmico de sistemas termodinâmicos em interação.

Tabela 4.8. Tabela de frequência das respostas às alternativas da questão 6.

alternativas		pré	pós	rep	total
a	88 °C	44,14	16,84	8,73	21,48
b	98 °C	21,62	19,90	23,81	21,48
c	110 °C	19,82	45,41	54,76	41,57
d	120 °C	14,41	16,84	12,70	15,01
8	Resposta inválida	0,00	0,00	0,00	0,00
9	Reposta faltando	0,00	1,020408	0,00	0,46
		100,00	100,00	100,00	100,00

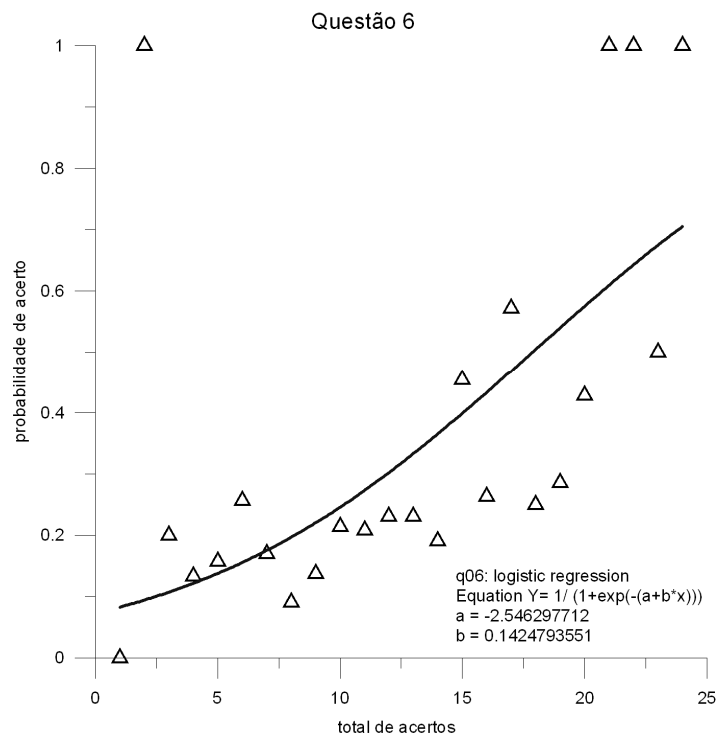


Figura 4.8. Curva Característica do Item da questão 6.

Pode-se também observar os resultados das questões 1 e 2, com suas respectivas tabelas de freqüências (Tabelas 4.9 e 4.10) e curvas características dos itens (Figuras 4.9 e 4.10). Essas questões foram selecionadas por estarem relacionadas a outro fenômeno de mudança de fase, a fusão.

Questão 1 - Qual é a temperatura mais provável dos cubos de gelo que se encontram armazenados no congelador de um freezer?

Tabela 4.9. Tabela de freqüência da questão 1.

alternativas		pré	pós	rep	total
a	-10°C	63,1	49,5	57,9	55,4
b	0°C	27,9	45,9	37,3	38,8
c	5°C	3,6	1,5	0	1,6
d	A temperatura depende do tamanho dos cubos de gelo	5,4	2,6	4,8	3,9
	Resposta inválida	0	0,5	0	0,2
	Resposta faltando	0	0	0	0
		100	100	100	100

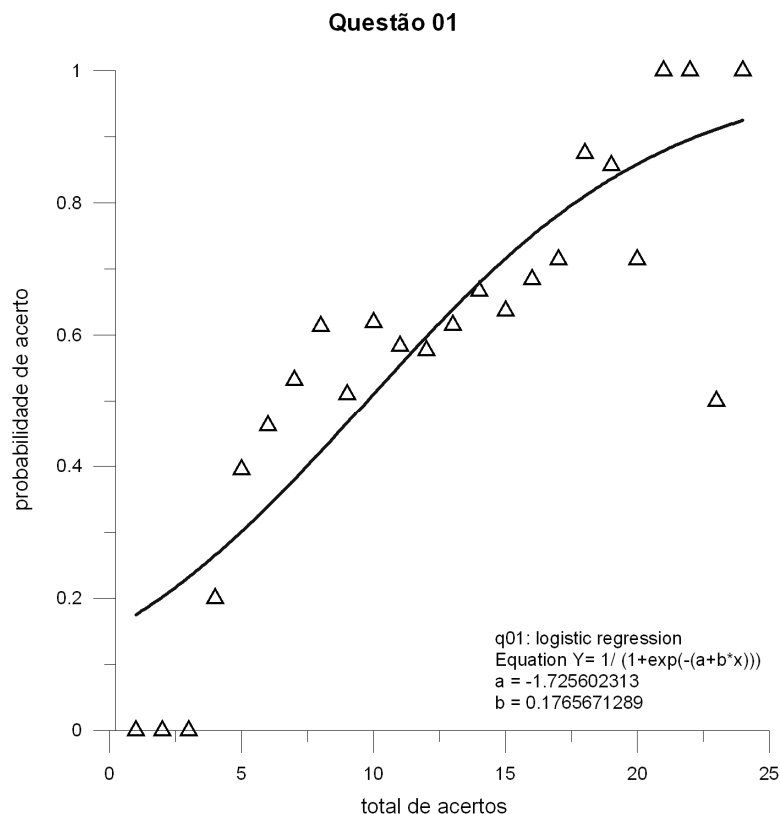


Figura 4.9. Curva Característica do Item da questão 1.

Questão 2: Francisco pega seis cubos de gelo no freezer e coloca quatro deles dentro de um copo com água, deixando os outros dois sobre o tampo de um balcão. Ele agita várias vezes o copo até que as pedras de gelo estejam muito pequenas e tenham parado de derreter. Qual é a temperatura mais provável da água nesta situação?

Tabela 4.10. Tabela de frequência da questão 2.

alternativas		pré	pós	rep	total
a	-10°C	5,4	1,0	2,4	2,5
b	0°C	26,1	41,3	54,0	41,1
c	5°C	41,4	38,3	28,6	36,3
d	10°C	27,0	17,9	15,1	19,4
8	Resposta inválida	0	1,5	0	0,69
9	Resposta faltando	0	0	0	0
		100	100	100	100

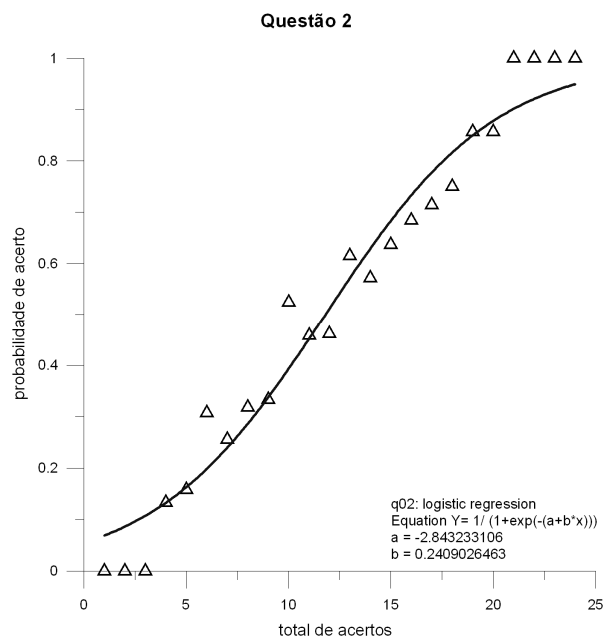


Figura 4.10. Curva Característica do Item da questão 1.

Na Questão 1, apesar da maioria dos alunos (55,4%), na média, marcarem a alternativa correta, observa-se um significativo número de estudantes (38,8%) que optam pelo distrator “b”, indicando, neste ponto, que se a água encontra-se no estado sólido (gelo), obrigatoriamente ela estará a 0°C. Um fator interessante de se notar é que a repetição do ensino formal do conteúdo não garantiu a melhoria do desempenho do aluno [Gonçalves Junior e Barroso 2011].

É também surpreendente, na questão 2, o número de alunos (55,7%, inclusive os que já passaram pelo ensino do tópico mais de uma vez) que assinalam um valor superior a 0°C para a temperatura da água em equilíbrio térmico com o gelo.

Numa análise conceitual do conjunto dessas 4 questões, percebe-se a não compreensão por parte dos alunos da lei zero da termodinâmica [Gonçalves Junior e Barroso 2011], ou seja, do conceito de equilíbrio térmico. Isto se evidencia tanto na diferenciação da temperatura da água em ebulição e do vapor formado logo acima dela, quanto nas diferentes temperaturas atribuídas ao gelo e água no processo de fusão.

Além disso, observa-se um elevado número de alunos que associam a mudança de fase a um valor pontual de temperatura e não à continuidade de troca de calor [Gonçalves Junior e Barroso 2011]. Essa concepção fica evidente quando o aluno não vê a possibilidade de existir água a 0°C, pois a essa

temperatura, para ele, a água se transforma em gelo, independente da mesma continuar a perder calor ou não. Isso é novamente evidenciado quando se trata da água em ebulição e o vapor formado nesse processo, isto é, como indicado nas questões 4 e 6, a partir do momento em que a água se transforma em vapor, a temperatura desse deve ser superior a 100°C.

Observando as Curvas Características do Item (ICC) para cada uma das questões selecionadas, podemos observar algumas características presentes na TRI através da construção das ICC's do modelo de dois parâmetros. A primeira delas é o índice de dificuldade que aparece em uma ordem crescente nas questões 1, 2, 4 e 6. Essa característica é observada diretamente das ICC's das questões, para o valor da nota correspondente a probabilidade de acerto de 50%, como apresentado na figura 4.11.

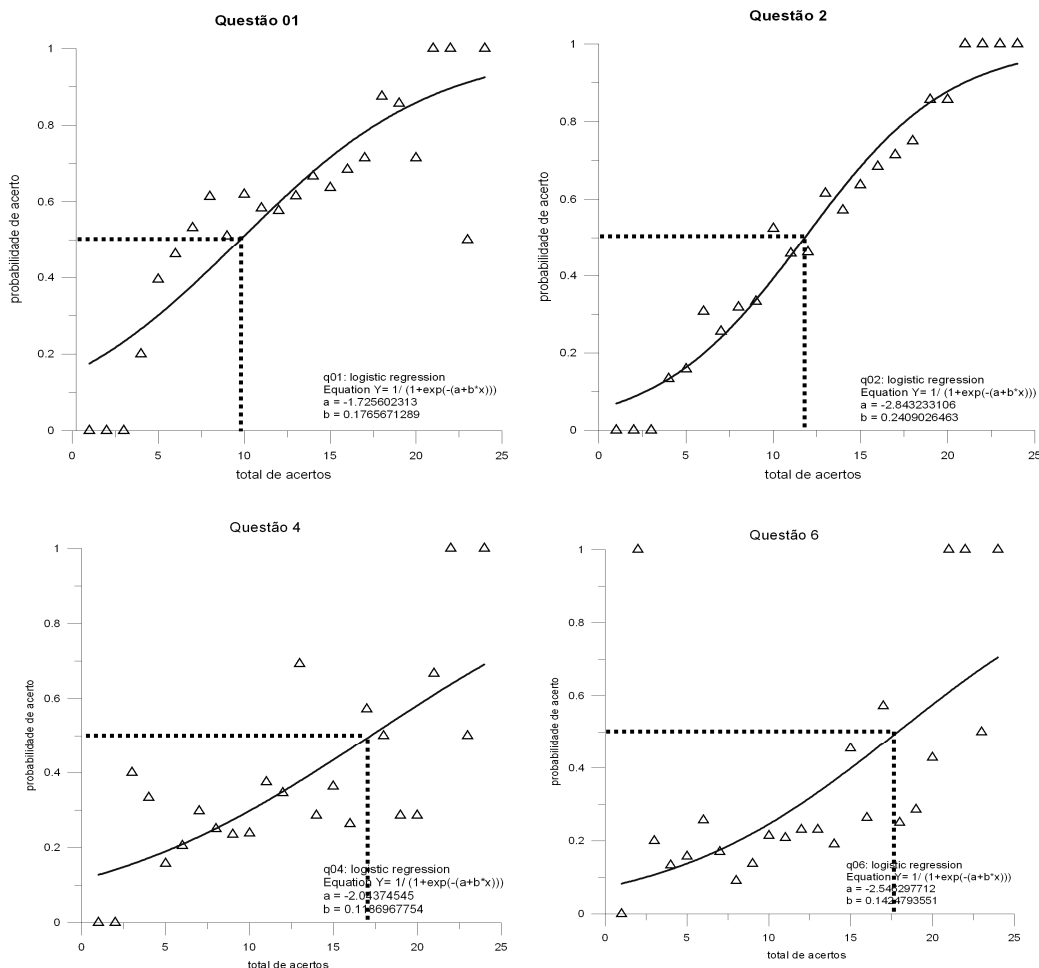


Figura 4.11. Identificação do índice de dificuldade da questão pelas ICC's

É possível, também, identificar nas questões 4 e 6, cujas curvas são mais adequadas para análises através de ICC's, o nível de discriminação da questão. Esta característica pode ser observada pela inclinação da curva na transição da baixa probabilidade de acerto para a alta probabilidade de acerto. Na figura 4.11, tem-se que a questão 6 possui maior índice de discriminação que a questão 4.

4.6- Considerações Finais

Pode-se concluir que o uso de questionários padronizados para avaliar as idéias dos estudantes e sua evolução é um recurso didático interessante. As possibilidades que um bom questionário, com uma análise cuidadosa, nos apresenta são as de abordar de forma diferenciada as questões que revelam as maiores dificuldades dos alunos.

No que se refere à aplicação do questionário no grupo estudado, observa-se uma dificuldade muito grande dos alunos em relação à compreensão do conceito de equilíbrio térmico. Aparentemente, os alunos entendem o congelamento da água como algo associado à sua temperatura, e não à perda de calor necessária para a mudança de estado. Em outras palavras, para esses estudantes é impossível encontrar água a 0°C. As diversas questões que contêm respostas que envolvem a perspectiva de existência de água em estado líquido a zero grau, exemplificadas no com as questões 1 e 2, indicam que os alunos reconhecem a existência de gelo a zero grau, mas não a existência de água a zero grau: a esta temperatura a água viraria necessariamente gelo.

O mesmo tipo de dificuldade é revelado em itens relativos à vaporização da água, como exemplificado nas questões 4 e 6. Nas respostas dadas a uma dessas questões predomina a idéia que a água ao se transformar em vapor tem a sua temperatura aumentada gradativamente, independente de estar recebendo calor da fonte geradora.

No caso do questionário proposto por Yeo e Zadnik [Yeo e Zadnik 2001], adaptado para aplicação em português, observa-se que algumas questões são de difícil avaliação sob a ótica da TRI. Isso porque ao traçarem-se as curvas

características dessas questões observa-se que elas não apresentam o comportamento esperado de uma questão formulada segundo as ideias da TRI, para a qual deve-se ter um único conceito presente em cada item: um comportamento crescente, no formato de um S para a probabilidade de acerto no item como função do desempenho do aluno.

Finalmente, a análise deste questionários tradicionais através da construção de curvas características do item se constituem uma rica ferramenta de análise dos itens aplicados.

CAPÍTULO 5 – Um Estudo do ENEM - Física

O ENEM, Exame Nacional do Ensino Médio, sofreu reformulações em 2009 principalmente em sua metodologia de aplicação e análise. De uma prova única, com avaliação de competências e habilidades de caráter pouco disciplinar e específico, utilizando a Teoria Clássica de Testes para a atribuição de escores, passou a um exame avaliando competências e habilidades associadas aos objetos de conhecimento, dentro de domínios cognitivos, com uma Matriz de Referência mais complexa e com a utilização da Teoria da Resposta ao Item.

A metodologia aplicada é descrita em notas técnicas disponíveis na página do INEP. Os escores foram normalizados (em 2009, para estudantes concluintes do ensino médio) de forma que a média de cada prova tivesse o valor 500 e desvio padrão 100.

Neste capítulo, apresentar-se-á um estudo qualitativo das questões de Física, parte integrante da prova de Ciências da Natureza e suas Tecnologias, a partir da reformulação do ENEM. Os microdados dessas provas não foram ainda disponibilizados pelo INEP, e a análise quantitativa desses dados será objeto de estudo posterior.

5.1 – Descrição Metodológica

Foram selecionadas as provas de Ciências da Natureza e suas Tecnologias do Exame Nacional do Ensino Médio a partir do ano de sua reformulação, em que se adotou a Teoria de Resposta ao Item como metodologia de análise.

As provas avaliadas são cinco: duas de 2009 (2009/1, a primeira versão, que não foi aplicada devido ao vazamento dentro da gráfica, e 2009/2, a segunda versão, aplicada aos estudantes), duas de 2010 (2010/1, a primeira aplicação, e 2010/2, a segunda aplicação, feita por defeitos de impressão da prova) e a de 2011.

Fez-se a restrição de análise às questões de Física dessas provas. Para fazer a seleção das questões, identificou-se como “questões de física” as que

abordam diretamente o conteúdo específico disciplinar de Física. No caso das questões que apresentaram traços interdisciplinares, multidisciplinares ou transdisciplinares, essa identificação foi feita de acordo com o conhecimento exigido para identificar a alternativa correta do item.

A partir dessa discriminação disciplinar, foram escolhidas as variáveis qualitativas cuja análise considerou-se relevante. A primeira delas foi reproduzir a classificação proposta pelo INEP para inclusão da questão no exame: a habilidade correspondente. Em seguida, estabeleceu-se um conjunto de parâmetros de interesse, associados ao tamanho dos textos em cada item, a existência ou não de objetos visuais (gráficos, tabelas, entre outros), o nível de contextualização do item, a exigência de conhecimentos específicos disciplinares de Física para resolução, a classificação da necessidade de cálculos quantitativos ou não para a resolução do item, além de aspectos associados à infra-estrutura da escola (especificamente, se o fato do aluno desenvolver atividades de laboratório torna mais fácil a solução das questões). Descreve-se a seguir com um pouco mais de detalhes cada uma dessas variáveis (e pretende-se, quando estiverem disponíveis os microdados das provas, fazer análise do efeito dessas variáveis nos resultados).

Após a discriminação disciplinar, fez-se a classificação das competências e habilidades exigidas em cada um dos itens selecionados. A partir de um método que vem sendo utilizado pelo INEP, buscou-se primeiramente a identificação da habilidade representativa desses itens. Em caso de dúvida entre duas ou mais habilidades para o mesmo item, identificou-se dentre essas habilidades qual estava apresentada, na Matriz de Referência, associada à competência mais adequada. Ainda restando dúvida entre duas habilidades dentro da mesma competência, procurou-se não repetir a habilidade, distribuindo os itens nas habilidades que compõem a referida competência.

Em seguida, foi feita a classificação dos conteúdos abordados nos itens analisados utilizando a classificação dos objetos de conhecimento associados à Matriz de Referência de Ciências da Natureza e suas Tecnologias (Apêndice 4).

A extensão dos itens foi avaliado usando como variável o número de linhas do texto-base e dos enunciados das questões. Esse número foi obtido considerando-se a formatação original da prova do ENEM, onde o caderno de

questão possui páginas divididas em duas colunas. Foram também levantadas a existência de imagens (I), tabelas (T) e gráficos (G), nesse texto básico, assim como a necessidade de utilizá-los para a resolução do item.

Quanto a contextualização, utilizou-se a classificação de Nentwig [Nentwig 2009], que estabelece que um item tem um alto nível de contextualização se no texto estão presentes as informações relevantes para sua solução, isto é, se a extração e processamento da informação contida no texto é necessária para a resolução do item; por outro lado, um baixo nível de contextualização está presente se a informação fornecida no texto não é essencial para responder ao item. Estabeleceu-se então a identificação do nível de contextualização do texto-base a partir de uma escala em que se atribui o valor “0” (zero) para a situação em que o texto constitui-se de um pré-texto [Machado 2010], ou seja, é completamente desnecessário para a solução do item. Se as informações necessárias para solução do item encontram-se entre 10% a 40% do número total de linhas do texto, o item foi classificado como um item de baixa contextualização, sendo atribuído a ele o grau “1”. No caso dessa porcentagem estar entre 50% a 70%, esse item será classificado como um item de média contextualização, sendo atribuído a ele o grau “2”. E, para finalizar, o item cujo texto-base possuía mais de 80% de seu corpo para resolução do mesmo foi considerado de alta contextualização, atribuindo-se grau “3”.

A exigência de se possuir conhecimento específico de física para a solução do item também foi avaliada. Para isso, quando a única exigência do item é a interpretação de informações contidas no texto, ou seja, a resposta está no item, não remetendo o candidato a conhecimentos específicos de física externos deste para sua resolução, a necessidade de conhecimento de conteúdo foi atribuído o termo “não”. Sendo exigida qualquer informação específica extra interpretação de texto, considerou-se que o item exigia “algum” nível de conhecimento físico e foi então atribuído a ele o termo “sim”. Paralelamente a essa classificação, levou-se em consideração a necessidade ou não da memorização de leis ou relações físicas expressas por equações matemáticas (conhecimento de “fórmulas”) para a solução do item, atribuindo-se o termo “não” (zero) quando não era necessária e o termo “sim” quando do oposto. Observa-se que a exigência de conhecimento de relações de proporcionalidade (como no

caso da determinação da resistência elétrica em função do material, seção transversal e comprimento do fio) foi considerada como exigências de “fórmula”, pois mesmo sem substituir valores e encontrar resultados numéricos, entende-se aqui a fórmula como a representação na linguagem matemática das proporcionalidades diretas e inversas existentes na descrição dos fenômenos físicos inquiridos.

Ainda dentro dessa linha, classificou-se os itens como quantitativos (qt), semi quantitativos (sq) e qualitativos (ql). Considerou-se aqui como quantitativos os itens que necessitam obrigatoriamente de cálculo para a obtenção da resposta da questão, que em geral, apresenta um valor ou intervalo numérico. Os semi quantitativos são aqueles itens que podem ser resolvidos pela análise de proporcionalidade, envolvendo obrigatoriamente o conhecimento do conceito físico e cuja análise se restrinja não a se encontrar um valor mas sim relações como “maior que”, “menor que”, “igual”, etc. Para finalizar, classificou-se como qualitativo aquele item que para sua solução necessita somente do conhecimento, compreensão ou análise do conceito físico, sem necessidade de utilização de raciocínio ou relação matemáticos.

Outro aspecto analisado foi a identificação de itens para os quais a existência de laboratório na escola ou a realização de atividades práticas poderiam, de alguma forma, auxiliar na solução do item proposto.

A seguir, apresenta-se os resultados da análise feita.

5.2- Resultados

5.2.1. O Número de Questões de Física

Na leitura das cinco provas mencionadas, observa-se uma regularidade do número de questões de física nas provas do ENEM durante os anos analisados. A prova de Ciências da Natureza contém 45 questões distribuídas em três áreas (Biologia, Física e Química); cerca de um terço da prova é destinada às questões de física. Na Figura 5.1, apresenta-se um gráfico de barras ilustrando o número de questões de Física para cada prova (2009/1 – a prova que não foi

aplicada; 2009/2 – a prova aplicada; 2010/1 – prova aplicada; 2010/2 – prova aplicada nos casos de problema com a impressão; 2011/1 – prova aplicada).

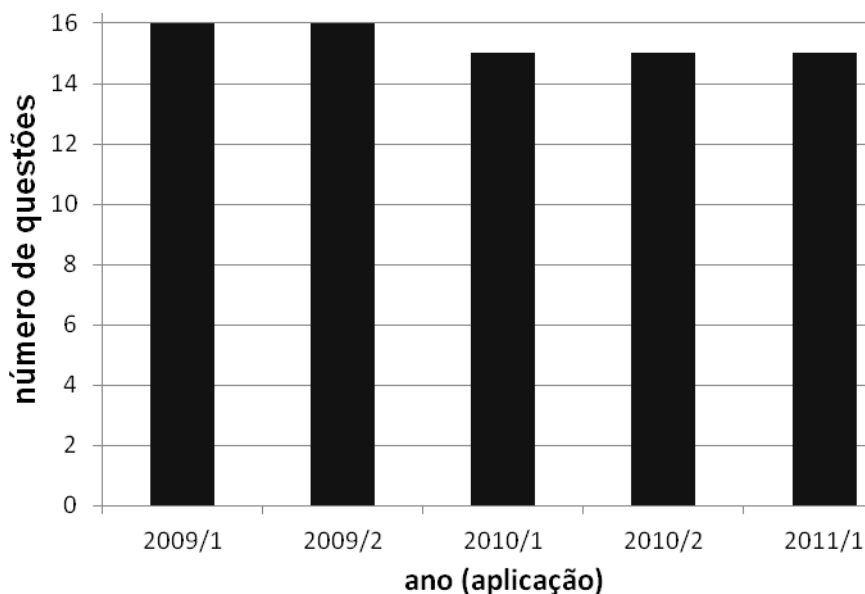


Figura 5.1. O número de questões de física nas provas do ENEM de 2009 a 2011

5.2.2. A Distribuição das Questões pelas Competências Previstas

Nas matrizes de referências do ENEM, a área das Ciências da Natureza é composta 8 competências distribuídas em 30 habilidades. As competências e suas respectivas habilidades, que a princípio permitem a construção dos itens identificados como da área curricular de Física, estão listadas na Tabela 5.1.

Tabela 5.1. Competências e Habilidades referentes à área de Física (segundo a Matriz de Referência do ENEM 2009)

<p>Competência de área 1 - Compreender as ciências naturais e as tecnologias a elas associadas como construções humanas, percebendo seus papéis nos processos de produção e no desenvolvimento econômico e social da humanidade.</p>	<p>H1 – Reconhecer características ou propriedades de fenômenos ondulatórios ou oscilatórios, relacionando-os a seus usos em diferentes contextos.</p>
	<p>H2 – Associar a solução de problemas de comunicação, transporte, saúde ou outro, com o correspondente desenvolvimento científico e tecnológico.</p>
	<p>H3 – Confrontar interpretações científicas com interpretações baseadas no senso comum, ao longo do tempo ou em diferentes culturas.</p>

<p>Competência de área 2 – Identificar a presença e aplicar as tecnologias associadas às ciências naturais em diferentes contextos.</p>	<p>H5 – Dimensionar circuitos ou dispositivos elétricos de uso cotidiano.</p>
	<p>H6 – Relacionar informações para compreender manuais de instalação ou utilização de aparelhos, ou sistemas tecnológicos de uso comum.</p>
	<p>H7 – Selecionar testes de controle, parâmetros ou critérios para a comparação de materiais e produtos, tendo em vista a defesa do consumidor, a saúde do trabalhador ou a qualidade de vida.</p>
<p>Competência de área 3 – Associar intervenções que resultam em degradação ou conservação ambiental a processos produtivos e sociais e a instrumentos ou ações científico-tecnológicos.</p>	<p>H8 – Identificar etapas em processos de obtenção, transformação, utilização ou reciclagem de recursos naturais, energéticos ou matérias-primas, considerando processos biológicos, químicos ou físicos neles envolvidos.</p>
	<p>H10 – Analisar perturbações ambientais, identificando fontes, transporte e(ou) destino dos poluentes ou prevendo efeitos em sistemas naturais, produtivos ou sociais.</p>
	<p>H12 – Avaliar impactos em ambientes naturais decorrentes de atividades sociais ou econômicas, considerando interesses contraditórios.</p>
<p>Competência de área 5 – Entender métodos e procedimentos próprios das ciências naturais e aplicá-los em diferentes contextos.</p>	<p>H17 – Relacionar informações apresentadas em diferentes formas de linguagem e representação usadas nas ciências físicas, químicas ou biológicas, como texto discursivo, gráficos, tabelas, relações matemáticas ou linguagem simbólica.</p>
	<p>H18 – Relacionar propriedades físicas, químicas ou biológicas de produtos, sistemas ou procedimentos tecnológicos às finalidades a que se destinam.</p>
	<p>H19 – Avaliar métodos, processos ou procedimentos das ciências naturais que contribuam para diagnosticar ou solucionar problemas de ordem social, econômica ou ambiental.</p>

Competência de área 6 – Apropriar-se de conhecimentos da física para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas.	H20 – Caracterizar causas ou efeitos dos movimentos de partículas, substâncias, objetos ou corpos celestes.
	H21 – Utilizar leis físicas e (ou) químicas para interpretar processos naturais ou tecnológicos inseridos no contexto da termodinâmica e(ou) do eletromagnetismo.
	H22 – Compreender fenômenos decorrentes da interação entre a radiação e a matéria em suas manifestações em processos naturais ou tecnológicos, ou em suas implicações biológicas, sociais, econômicas ou ambientais.
	H23 – Avaliar possibilidades de geração, uso ou transformação de energia em ambientes específicos, considerando implicações éticas, ambientais, sociais e/ou econômicas.

Analisando-se as provas do ENEM de 2009 a 2011 e levando-se em conta os critérios pré-estabelecidos, as competências referentes às questões de física apresentaram a distribuição apresentada na Figura 5.2.

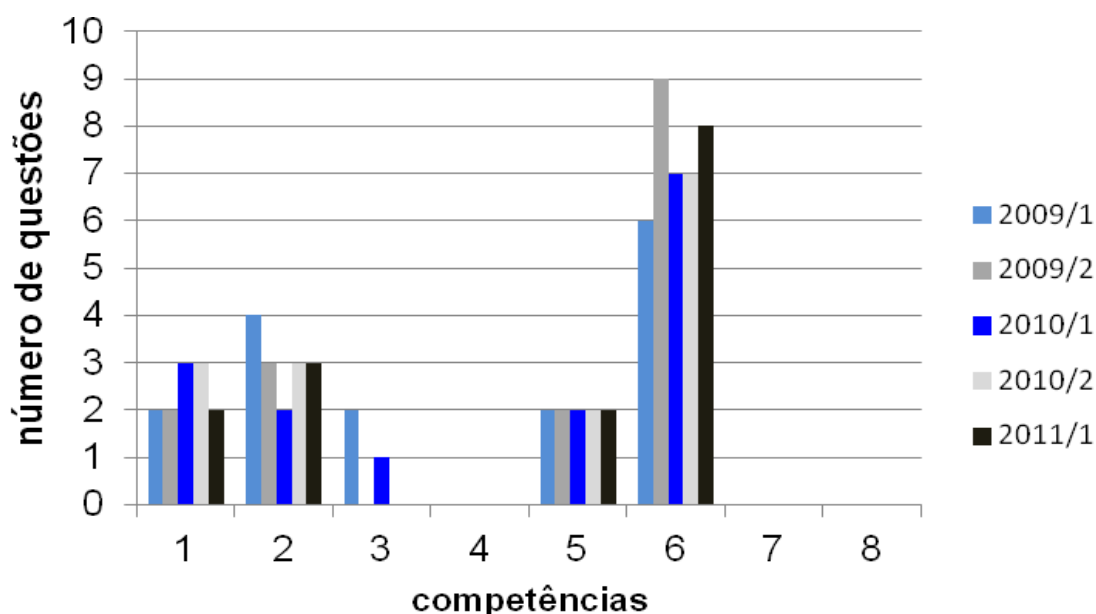


Figura 5.2. Distribuição das competências na prova do ENEM por aplicação.

Analisando a Figura 5.2, observa-se que, excetuando-se as competências 3 e 4, que contemplam as áreas de biologia e de química, as outras competências listadas na Tabela 5.1 foram solicitadas em todas as provas do período analisado neste trabalho. No entanto, existe uma incidência maior de itens na competência 6. Isto certamente pode ser atribuído ao fato de que essa competência se refere especificamente aos conteúdos de física, o que facilita a elaboração de itens dessa disciplina.

5.2.3. As habilidades das questões

Dando continuidade à análise, organizou-se a distribuição das habilidades presentes nos itens das provas dentro de suas competências; o resultado está apresentado nas Tabelas 5.2 a 5.6.

Tabela 5.2. Distribuição das habilidades na competência de área 1

Competência	Habilidade	2009/1	2009/2	2010/1	2010/2	2011/1
1	1	2	0	2	2	2
	2	0	1	0	0	0
	3	0	1	1	1	0

Tabela 5.3. Distribuição das habilidades na competência de área 2

Competência	Habilidade	2009/1	2009/2	2010/1	2010/2	2011/1
2	5	1	2	1	2	1
	6	2	1	1	0	1
	7	1	0	0	1	1

Tabela 5.4. Distribuição das habilidades na competência de área 3

Competência	Habilidade	2009/1	2009/2	2010/1	2010/2	2011/1
3	8	0	0	1	0	0
	10	0	0	0	0	0
	12	0	0	0	0	0

Tabela 5.5. Distribuição das habilidades na competência de área 5

Competência	Habilidade	2009/1	2009/2	2010/1	2010/2	2011/1
5	17	1	1	1	0	2
	18	0	1	1	2	0
	18	1	0	0	0	0

Tabela 5.6. Distribuição das habilidades na competência de área 6

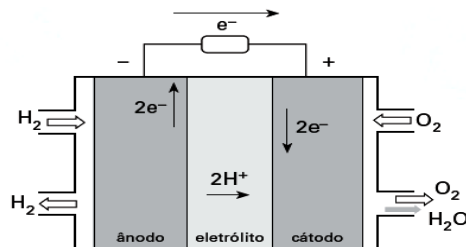
Competência	Habilidade	2009/1	2009/2	2010/1	2010/2	2011/1
6	20	1	2	1	2	3
	21	1	4	2	2	1
	22	2	2	1	1	1
	23	2	1	3	2	3

A observação dessas Tabelas 5.2 a 5.6 permite constatar que as 16 habilidades de Ciências Naturais, que contemplam a área de conhecimento da Física, não são necessariamente solicitadas em todas as provas nas questões desta disciplina. Isto pode ser explicado, no entanto, em decorrência dos itens elaborados para a prova terem sua construção baseada nas habilidades da matriz do ENEM e não no conteúdo específico de uma disciplina. Esta condição permite que uma habilidade que não esteja presente nos itens de Física possa ser contemplada, em alguns casos, pelas áreas de Biologia ou de Química.

Como contraponto, observa-se que na competência de área 6, que é a única destinada exclusivamente à disciplina de Física, encontra-se numa mesma prova mais de um item contemplando a mesma habilidade.

No entanto, necessita-se ressaltar que ao se fazer as classificações dos itens das provas, em suas respectivas habilidades, existem vários itens que se encaixam em mais de uma habilidade. Observe, por exemplo, a questão número 63 utilizada na primeira aplicação da prova de 2010:

O crescimento da produção de energia elétrica ao longo do tempo tem influenciado decisivamente o progresso da humanidade, mas também tem criado uma séria preocupação: o prejuízo ao meio ambiente. Nos próximos anos, uma nova tecnologia de geração de energia elétrica deverá ganhar espaço: as células a combustível hidrogênio/oxigênio.



VILLULLAS, H. M; TICIANELLI, E. A; GONZÁLEZ, E. R. *Química Na Nova Escola*. Nº 15, maio 2002.

Com base no texto e na figura, a produção de energia elétrica por meio da célula a combustível hidrogênio/oxigênio diferencia-se dos processos convencionais porque

- A) transforma energia química em energia elétrica, sem causar danos ao meio ambiente, porque o principal subproduto formado é a água.
- B) converte a energia química contida nas moléculas dos componentes em energia térmica, sem que ocorra a produção de gases poluentes nocivos ao meio ambiente.
- C) transforma energia química em energia elétrica, porém emite gases poluentes da mesma forma que a produção de energia a partir dos combustíveis fósseis.
- D) converte energia elétrica proveniente dos combustíveis fósseis em energia química, retendo os gases poluentes produzidos no processo sem alterar a qualidade do meio ambiente.
- E) converte a energia potencial acumulada nas moléculas de água contidas no sistema em energia química, sem que ocorra a produção de gases poluentes nocivos ao meio ambiente.

Figura 5.3. Questão 63 da 1ª aplicação do ENEM 2010.

Ao se buscar na Matriz de Referências do Enem a habilidade contemplada por esse item, encontram-se as seguintes possibilidades [Matriz de Referência 2009, apresentada no Anexo 2]:

H8 – Identificar etapas em processos de obtenção, transformação, utilização ou reciclagem de recursos naturais, energéticos ou matérias-primas, considerando processos biológicos, químicos ou físicos neles envolvidos.

H19 – Avaliar métodos, processos ou procedimentos das ciências naturais que contribuam para diagnosticar ou solucionar problemas de ordem social, econômica ou ambiental.

H23 – Avaliar possibilidades de geração, uso ou transformação de energia em ambientes específicos, considerando implicações éticas, ambientais, sociais e/ou econômicas.

H26 – Avaliar implicações sociais, ambientais e/ou econômicas na produção ou no consumo de recursos energéticos ou minerais, identificando transformações químicas ou de energia envolvidas nesses processos.

Qualquer uma das 4 habilidades satisfazem esse item. Isso se torna um problema quando se leva em conta que a Teoria de Resposta ao Item, utilizada para analisar e gerar as notas da prova do ENEM, pressupõe que o item construído deve contemplar somente uma habilidade para que o mesmo possa ser utilizado na análise. Este fato aponta para um problema na construção da Matriz de Referência do ENEM, que permite que sejam construídos itens que pertençam a mais de uma habilidade. Observa-se que esse fato ocorre com vários outros itens, como mostrado na Tabela 5.7.

Tabela 5.7. Distribuição das habilidades na competência de área.

Ano/Aplicação	Número do Item	Habilidades Contempladas
2009/1	1	9 ou 21
	9	9, 12 ou 21
	28	22 ou 23
	31	17 ou 21
	33	17, 19 ou 23
2009/2	14	5, 19 ou 23
	17	18 ou 20
	18	5 ou 6

	19	5 ou 17
	24	17 ou 21
	31	1 ou 2
	32	1 ou 22
	35	17, 21 ou 22
2010/1	47	1 ou 22
	48	5 ou 17
	50	3 ou 21
	52	6, 18 ou 23
	54	1 ou 22
	59	21 ou 23
	63	8, 19, 23 ou 26
	70	5, 6, 17 ou 18
	84	1 ou 18
2010/2	48	17 ou 23
	57	5 ou 18
	59	5 ou 17
	64	17 ou 20
	65	18 ou 20
	67	1, 18 ou 22
	81	18 ou 21
	84	5 ou 18
	85	1 ou 22
	86	18 ou 23
	90	6, 7, 17 ou 18
2011/1	46	17 ou 20
	60	5, 6 ou 17
	63	1, 17 ou 22
	67	17 ou 22
	74	1, 18 ou 22
	77	17 ou 20

Nesse ponto, faz-se um parêntese para que se possa analisar as habilidades 17 e 18, que são assim enunciadas na Matriz de Referência do ENEM:

H17 – Relacionar informações apresentadas em diferentes formas de linguagem e representação usadas nas ciências físicas, químicas ou biológicas, como texto discursivo, gráficos, tabelas, relações matemáticas ou linguagem simbólica.

H18 – Relacionar propriedades físicas, químicas ou biológicas de produtos, sistemas ou procedimentos tecnológicos às finalidades a que se destinam.

As duas redações são generalistas e amplas, de forma que grande parte das questões podem ser classificadas em uma das duas. Isso fica evidenciado ao se consultar a Tabela 5.7, onde estas duas habilidades são as que mais dividem itens com as demais habilidades indicadas. Reforça-se aqui que mais uma vez é fragilizada a estrutura da TRI que exige seus itens sejam unidimensionais, ou seja, contemplem uma única habilidade.

5.2.4. Os Objetos de Conhecimento

Os objetos de conhecimento associados à Matriz de Referência de Ciências da Natureza e suas Tecnologias são divididos (de acordo com a classificação apresentada no Apêndice 4) em 7 grandes áreas:

1. Conhecimentos básicos e fundamentais;
2. O movimento, o equilíbrio e a descoberta de leis físicas;
3. Energia, trabalho e potência;
4. A mecânica e o funcionamento do universo;
5. Fenômenos elétricos e magnéticos;
6. Oscilações, ondas, óptica e radiação;
7. O calor e fenômenos térmicos.

A partir desta classificação, foi construído o gráfico do número de itens por área, em cada ano de aplicação da prova, apresentado nas Figuras 5.4 a 5.8.

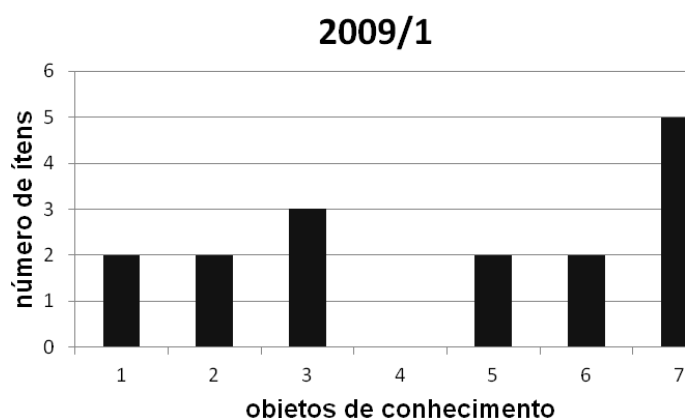


Figura 5.4. Número de itens por objeto de conhecimento - prova ano 2009/1.

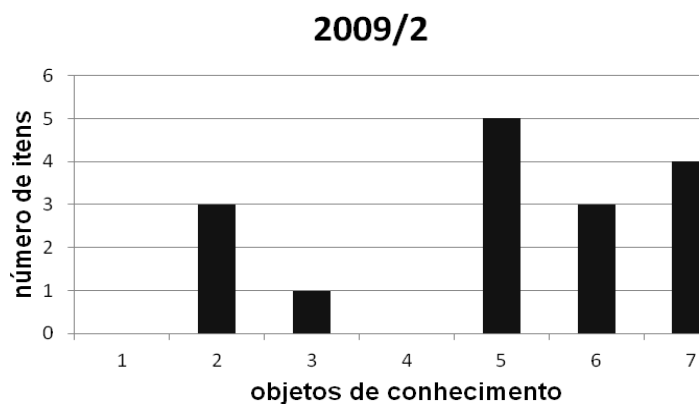


Figura 5.5. Número de itens por objeto de conhecimento - prova ano 2009/2.

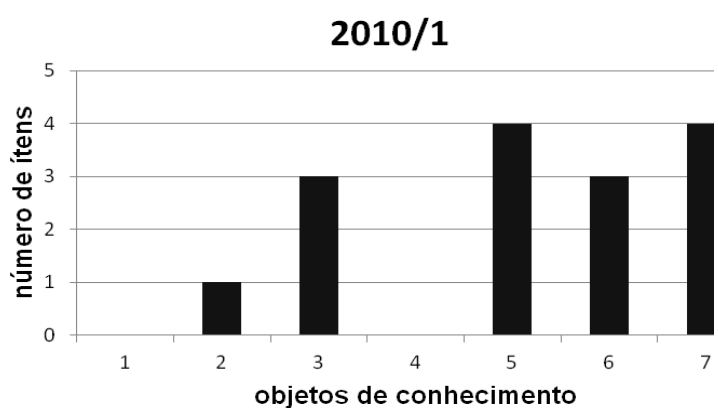


Figura 5.6. Número de itens por objeto de conhecimento - prova ano 2010/1.

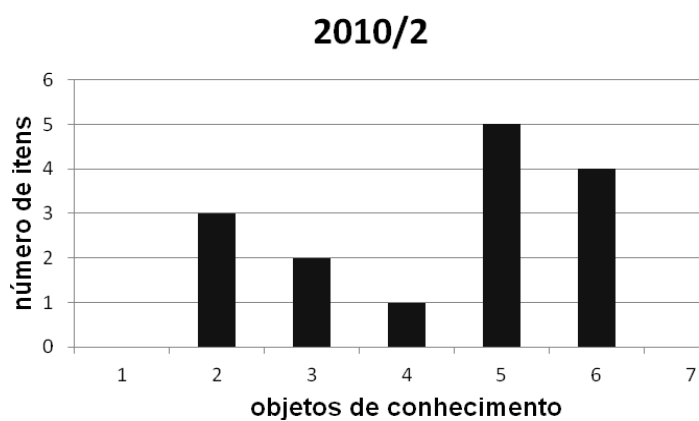


Figura 5.7. Número de itens por objeto de conhecimento - prova ano 2010/2.

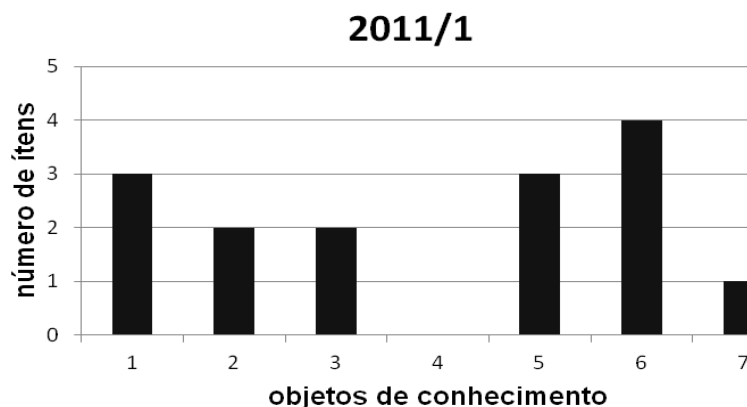


Figura 5.8. Número de itens por objeto de conhecimento - prova ano 2011/1.

O primeiro fato que se pode observar é que existe uma regularidade na distribuição dos conteúdos, sendo que os objetos de conhecimento das áreas 5, 6 e 7 são privilegiados nessa distribuição, compondo 56% da prova em 2009/1, 75% em 2009/2, 74% em 2010/1, 60% em 2010/2 e 54% em 2011.

No entanto, quando se volta a atenção para os objetos relacionados aos conteúdos de mecânica (áreas 2, 3 e 4), que compõem relativamente a maior parte da grade do ensino médio, observa-se um pequeno percentual dedicado a essa área, sendo de 31,5% na prova de 2009/1, 25% em 2009/2, 26% em 2010/1, 40% em 2010/2 e de 26% em 2011.

A análise da presença do conteúdo de mecânica, quando observada com mais cuidado, revela que a área 2, na qual encontra-se todo conteúdo de cinemática, dinâmica, estática da partícula e do corpo rígido, conservação da quantidade de movimento e colisões e por fim mecânica dos fluidos, não incluindo somente a parte de trabalho e energia mecânica (que compõem a área 3), foi contemplada com 12,5% dos itens de física na prova de 2009/1, 19% em 2009/2, 6% em 2010/1, 20% em 2010/2 e finalmente 13% em 2011. Observa-se aqui, então, que a maior parte do conteúdo de mecânica tem sido indiscriminadamente desprivilegiado nas provas do ENEM. Essa distorção pode ser em decorrência de haver uma única habilidade que contemple especificamente essa área, a habilidade 20, que consiste em “Caracterizar causas ou efeitos dos movimentos de partículas, substâncias, objetos ou corpos celestes”.

Ainda cabe neste ponto, ressaltar que os itens classificados nos objetos de conhecimento de área 3, quando vistos com maior detalhamento, revelam que na sua maioria compreendem questões gerais de transformação de energia, não contemplando de forma específica o conhecimento de energia e trabalho mecânicos, como indicado na Tabela 5.8.

Tabela 5.8. Objetos de conhecimentos específicos da área 3.

Ano/Aplicação	Número da Questão	Objeto de Conhecimento Específico
2009/1	19	3.1
	33	3.1
	42	3.1
2009/2	20	3.1
2010/1	56	3.1
	63	3.1
	89	3.1
2010/2	48	3.1
	86	3.1
2011/1	80	3.1
	86	3.2

Praticamente todas as questões referentes à área 3 contemplam o conteúdo específico 3.1, ou seja, “conceituação de trabalho, energia e potência”. Observe a questão 48 da prova de Ciências da Natureza aplicada em 2010/2:

No nosso dia a dia deparamo-nos com muitas tarefas pequenas e problemas que demandam pouca energia para serem resolvidos e, por isso, não consideramos a eficiência energética de nossas ações. No global, isso significa desperdiçar muito calor que poderia ainda ser usado como fonte de energia para outros processos. Em ambientes industriais, esse reaproveitamento é feito por um processo chamado de cogeração. A figura a seguir ilustra um exemplo de cogeração na produção de energia elétrica.

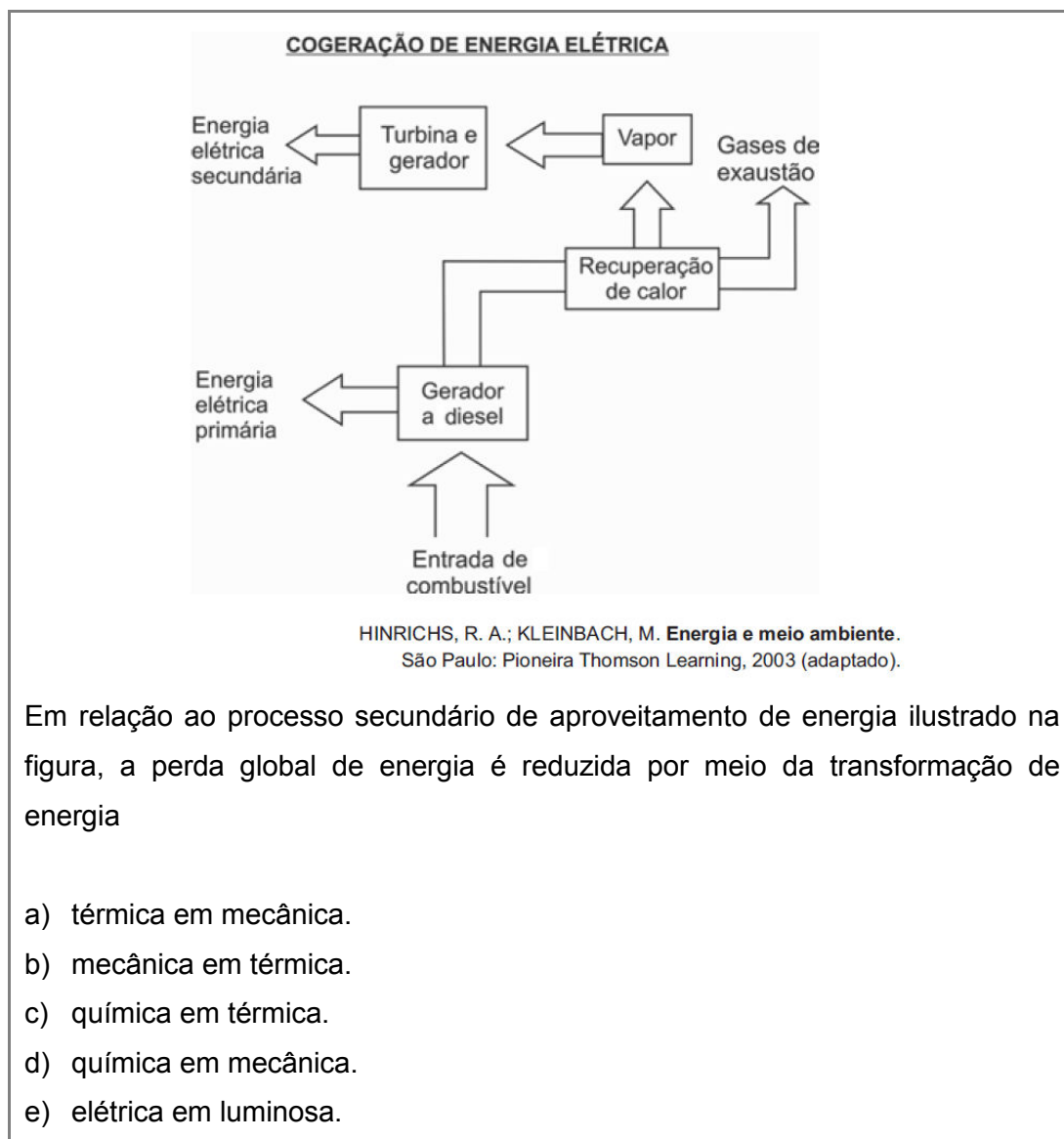


Figura 5.9. Questão 48 da 2ª aplicação do ENEM 2010.

Esse tipo de abordagem do conhecimento compreendido no item tornou-se comum em virtude de sua aplicação nas provas do ENEM, e muito pouco representam a área de conhecimento do conteúdo de Trabalho e Energia Mecânica.

Essa discussão da distribuição dos conteúdos nos itens de física torna-se fundamental a partir do momento em que o ENEM, a partir de 2011, passa a ser a única forma de acesso a maior parte das universidades públicas brasileiras e, portanto, passa a ser referência e conseqüentemente parâmetro para a determinação do currículo aplicado e ensinado nas escolas do Ensino Médio. O que se pode deduzir dos dados colhidos, nas provas até aqui aplicadas, é que a

Mecânica, que até então se constitui um dos pilares do Ensino de Física nas escolas do Ensino Médio, passa a ser um conhecimento menos exigido e como consequência, menos ensinado. Apresenta-se, nesse ponto, com essa análise, mais um indício dos problemas existentes na construção da Matriz de Referência, que de forma indiscriminada e sem apresentar um estudo sobre alteração dos currículos do Ensino Médio no Brasil, passa a promover mudanças tão profundas no mesmo.

5.2.5. A Extensão dos Itens

O grande número de itens existentes nas provas do ENEM (180 itens divididos em duas provas) está relacionado às características das avaliações que utilizam a TRI como metodologia. No entanto, é necessário, como em qualquer avaliação, que os candidatos façam de forma consciente e séria os itens propostos. Nesse contexto, encontra-se outro ponto a ser discutido: a extensão do texto-base dos itens da prova do ENEM. Como se é sabido, os candidatos ao realizar a prova dispõem, em média, de 3 minutos para a solução de cada item. No entanto, os textos-base dos itens propostos possuem, de forma geral, grande extensão, tornando quase impossível que o aluno consiga fazê-los no tempo estipulado para a prova. Esta característica pode ser percebida quando se conhece a média aritmética do número de linhas do texto base das questões.

Tabela 5.9. Média do número de linhas do texto-base dos itens do ENEM por ano/aplicação

Ano/Aplicação	Extensão do Texto-Base (média)	Desvio Padrão
2009/1	11	5
2009/2	10	3
2010/1	9	3
2010/2	9	5
2011/1	11	4

Para ratificar as informações da tabela, construíram-se os diagramas onde, por ano/aplicação, estão representadas as questões e seus respectivos número de linhas de texto-base.

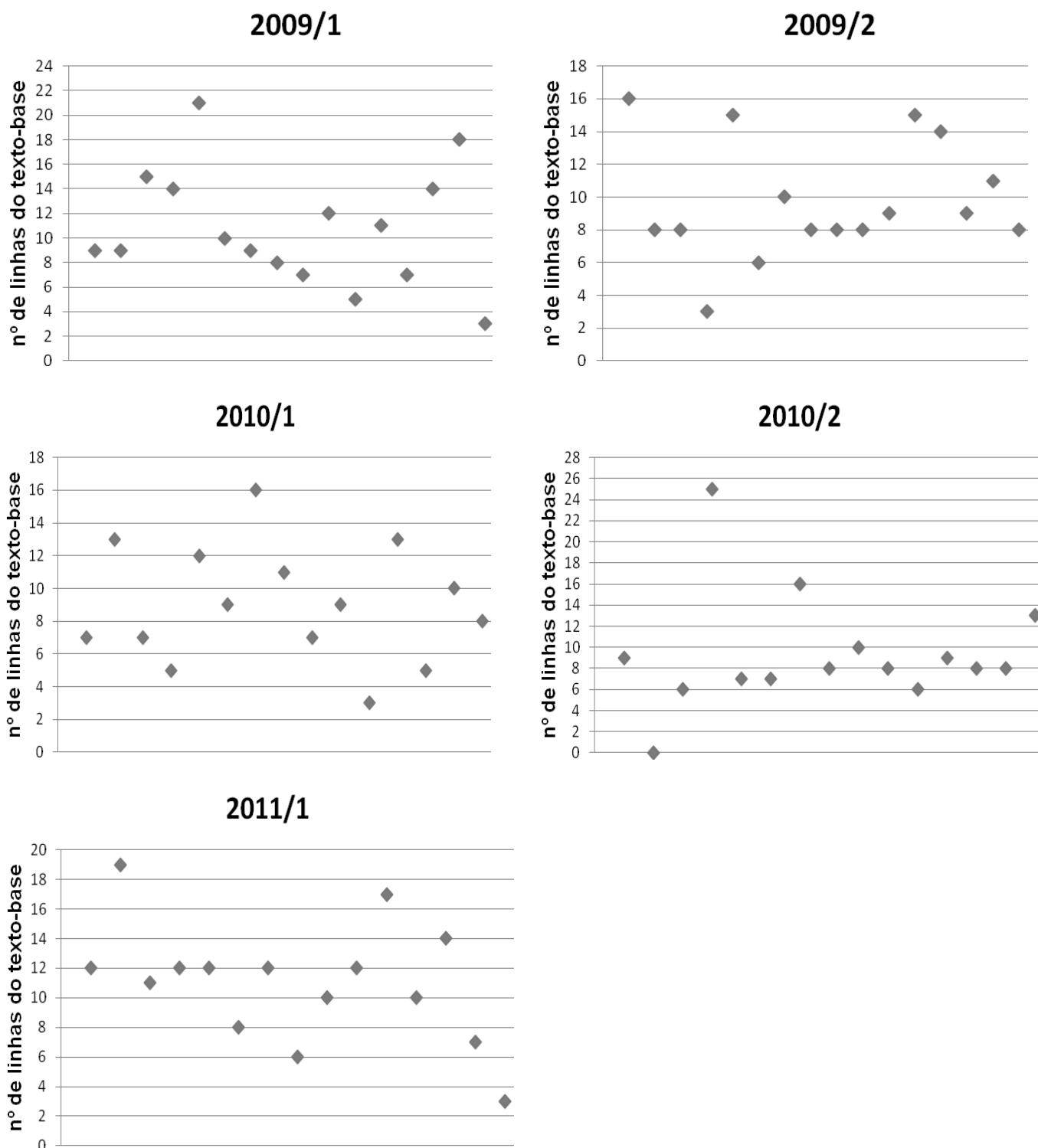


Figura 5.10. Número de linhas do texto-base dos itens do ENEM por ano/aplicação.

Pode-se observar na figura 5.10 que o número de linhas do texto-base é consideravelmente elevado para um tempo de 3 minutos destinados à resolução

do item, sem contar que alguns dos itens ainda exigem a leitura, interpretação ou análise de gráficos, imagens ou tabelas.

Outro detalhe que não se pode ser esquecido é o fato de que o candidato possui o enunciado (pergunta do item) e as alternativas para ler, compreender e responder o item. No caso do enunciado, como se pode observar na tabela 5.10, das médias do número de linhas do enunciado por ano/aplicação, não pode simplesmente ser desprezado quando se fala de um tempo tão curto para resolução.

Tabela 5.10. Média do número de linhas do enunciado dos itens do ENEM por ano/aplicação.

Ano/Aplicação	Extensão do Enunciado (média)	Desvio Padrão
2009/1	4	3
2009/2	5	3
2010/1	3	1
2010/2	3	1
2011/1	3	1

Pode-se concluir, portanto, que há necessidade de se elaborar itens menos extensos para que se possa garantir a qualidade das respostas da prova no que se diz respeito ao tempo necessário para sua execução.

5.2.6. A Utilização de Imagens, Gráficos e Tabelas no Texto-Base

Continuando o trabalho, fez-se um estudo da presença de imagens (I), gráficos (G) e tabelas (T) nas provas, bem como a classificação se as mesmos eram necessários para a resolução do item. A Tabela 5.11 apresenta os dados obtidos.

Tabela 5.11. Itens que apresentam imagens, gráficos e tabelas.

Ano/Aplicação	Número da Questão	Tipo de Linguagem	É necessário para resolução do item?
2009/1	17	Imagem	sim
	23	Tabela + Gráfico	sim
	28	Imagem	sim
	31	Imagem	sim
	32	Imagem	não
	33	Tabela	sim
	45	Tabela	Sim
2009/2	14	Imagem	sim
	18	Tabela	sim

	19	Tabela + Tabela + Imagem	sim
	20	Imagem	sim
	24	Gráfico	sim
	27	Imagem	não
	32	Imagem	sim
	35	Imagem	sim
	39	Imagem	não
	45	Imagem	Sim
2010/1	48	Tabela	sim
	54	Imagem	sim
	56	Imagem	sim
	63	Imagem	sim
	68	Imagem	sim
	70	Tabela	Sim
2010/2	48	Gráfico	sim
	49	Imagem	sim
	59	Tabeça	sim
	64	Gráfico	sim
	65	Imagem + Imagem	sim
	81	Imagem	sim
	84	Imagem	não
	90	Imagem	sim
2011/1	46	Imagem	sim
	54	Tabela	sim
	60	Tabela	sim
	63	Imagem + Imagem	sim
	67	Gráfico	sim
	70	Imagem	sim
	73	Imagem	sim
	77	Tabela	sim
	78	Imagem	sim
	84	Imagem	sim
	86	Imagem	sim

Da Tabela 5.11 pode-se inferir que não existe uma regularidade no número de imagens, tabelas ou gráficos utilizados nos itens e, a não ser por algumas exceções, a utilização desses elementos é feita somente quando eles são essenciais para a resolução do item em questão. Este fato indica uma preocupação, por parte do INEP, na direção de não se utilizar imagens como mera alegoria, as usando somente quando se constituem parte essencial do texto-base.

5.2.7. Nível de Contextualização dos Itens

Outra análise feita foi a classificação do nível de contextualização dos itens de física, sendo que o resultado obtido pode ser observado na figura

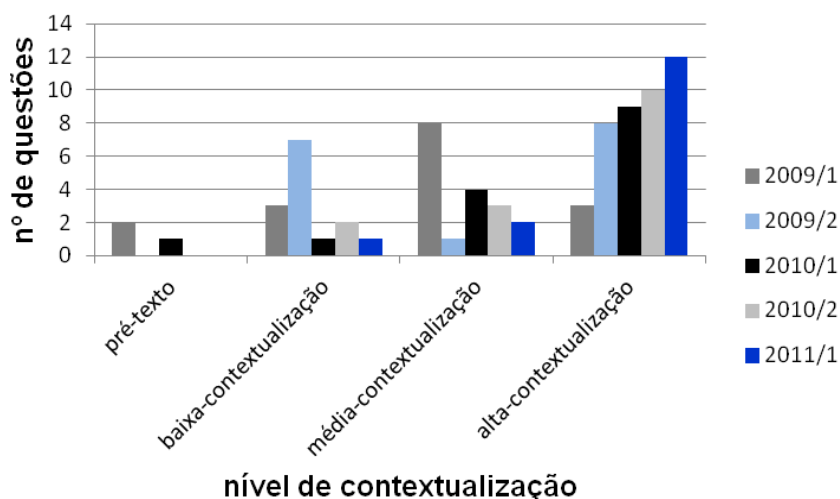


Figura 5.11. Nível de contextualização do itens por ano/aplicação.

Como se pode constatar, dentro dos critérios pré-estabelecidos para essa classificação [Nentwig 2009], o nível de contextualização dos itens melhorou significativamente, de forma que, no ano de 2011, tem-se 80% dos itens de física com um alto nível de contextualização. Pode-se observar essa evolução no crescimento do número de itens com alta contextualização, representado no último trecho do gráfico.

5.2.8. Exigência de Conhecimentos Específicos de Física

Levando-se em consideração que o ENEM, após sua reformulação, passou a ser o único meio de ingresso em muitas universidades públicas e privadas, faz-se importante que ele avalie, em algum nível, o conhecimento dos alunos em áreas específicas como, por exemplo, a de Física

No entanto, apesar da melhora no nível de contextualização dos itens, ainda existem alguns deles que não exigem conhecimento específico de física, solicitando somente interpretação de texto e, algumas vezes, da noção de

proporção (regra de três) para sua solução. Isto pode ser observado, por exemplo, na questão 63 do ENEM 2011, apresentada na figura 5.12.

QUESTÃO 63

Para que uma substância seja colorida ela deve absorver luz na região do visível. Quando uma amostra absorve luz visível, a cor que percebemos é a soma das cores restantes que são refletidas ou transmitidas pelo objeto. A Figura 1 mostra o espectro de absorção para uma substância e é possível observar que há um comprimento de onda em que a intensidade de absorção é máxima. Um observador pode prever a cor dessa substância pelo uso da roda de cores (Figura 2): o comprimento de onda correspondente à cor do objeto é encontrado no lado oposto ao comprimento de onda da absorção máxima.

Figura 1

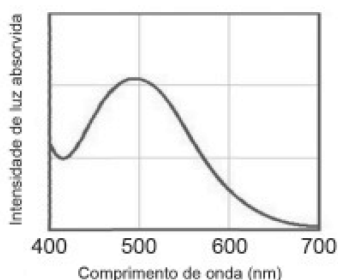
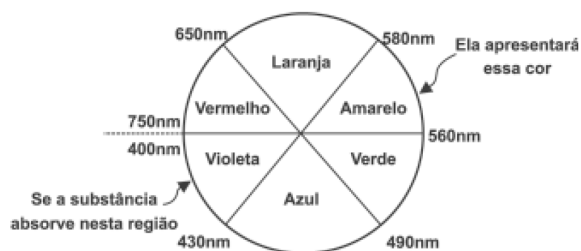


Figura 2



Brown, T. *Química a Ciência Central*. 2005 (adaptado).

Qual a cor da substância que deu origem ao espectro da Figura 1?

- A Azul.
- B Verde.
- C Violeta.
- D Laranja.
- E Vermelho.

Figura 5.12. Questão 63 da aplicação do ENEM 2011.

Observa-se que a leitura e a interpretação do texto-base são suficientes para a resolução desta questão; logo, esse item não representa uma habilidade exclusiva da área das Ciências da Natureza, podendo, por exemplo, ser inserida na prova de Matemática e suas Tecnologias, na competência 6, habilidade 24, 25 ou 26 que consistem em:

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

Habilidade 24 – Utilizar informações expressas em gráficos ou tabelas para fazer inferências.

Habilidade 25 – Resolver problema com dados apresentados em tabelas ou gráficos.

Habilidade 26 – Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.

Na Figura 5.13, o diagrama indica que não há preocupação com esse tipo de ocorrência, pois, após uma aparente tendência de diminuição desse tipo de item entre anos de 2009 e 2010, verifica-se a volta do crescimento dos mesmos em 2011, constituindo-se de 33% dos itens que podem ser classificados no conhecimento específico de física.

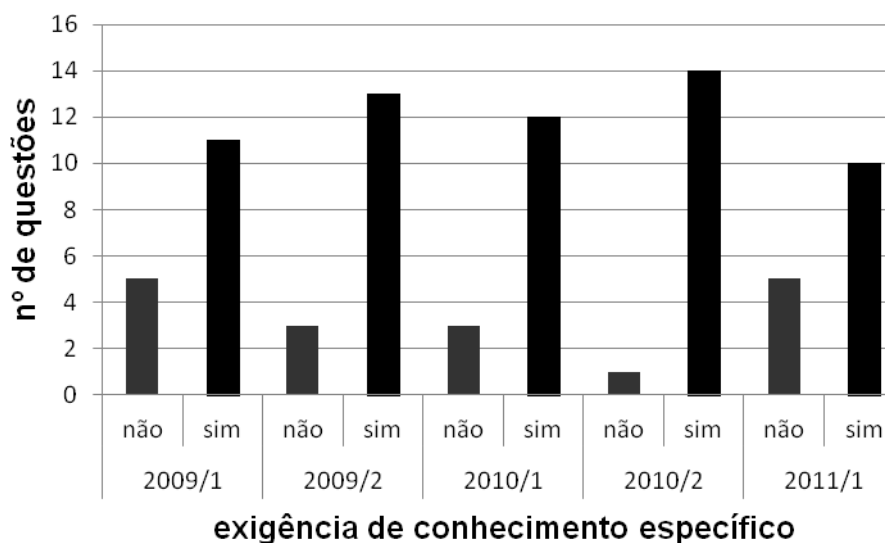


Figura 5.13. Exigência de cobrança de conhecimento específico de física na prova do ENEM.

Na análise dos aspectos relacionados à existência de cobrança específica de conteúdo de física no exame, é necessário comentar o fato da apresentação, nas provas, de questões com erros conceituais. Isso aconteceu, por exemplo, com o item 54 do ENEM 2010/1 e o item 46 do ENEM 2011 (numeração em referência à prova azul). Dadas as proporções que o ENEM tomou a partir de sua modificação, indica-se aqui a necessidade de uma revisão de conteúdo mais cuidadosa das provas.

5.2.9. Natureza dos Itens do ENEM

Como pode ser observado na Figura 5.14., dentro das questões que exigem algum conhecimento específico de física, existe uma pequena parte (25% \pm 5%) que solicita do candidato o conhecimento de relações matemáticas, como definido anteriormente, para a resolução do item.

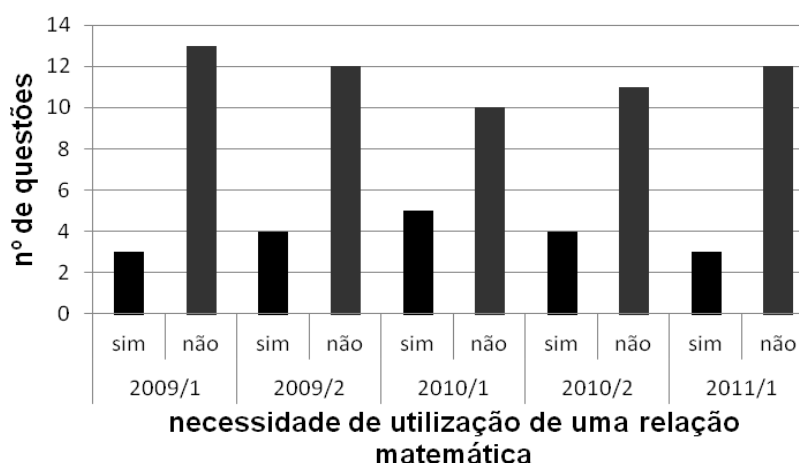


Figura 5.14. Exigência de cobrança de relações matemáticas na prova do ENEM.

O resultado observado na Figura 5.14 confirma a observação feita a partir Figura 5.15, onde se observa que a prova de física do ENEM é predominantemente qualitativa, privilegiando muito pouco a resolução de problemas numéricos.

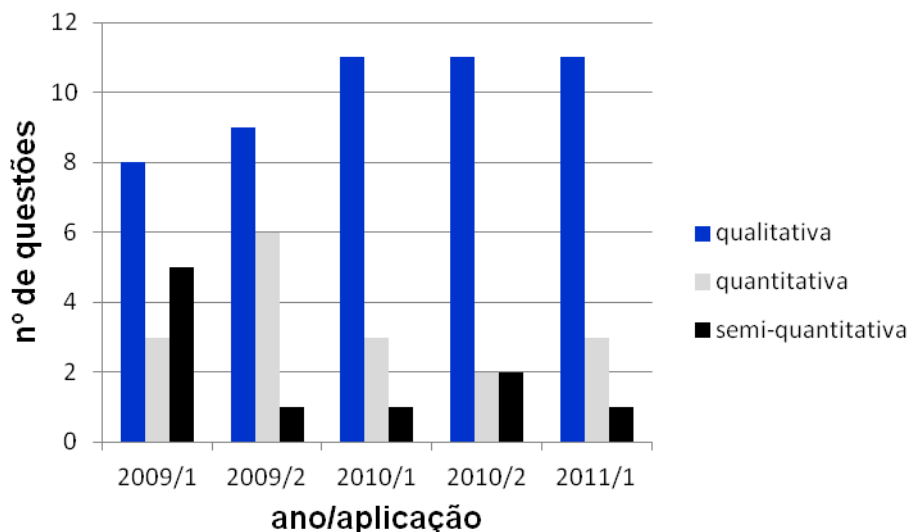


Figura 5.15. Exigência de cobrança de relações matemáticas na prova do ENEM.

5.3. Considerações finais

Neste Capítulo, apresentamos uma análise da componente disciplinar Física das provas de Ciências da Natureza do ENEM utilizadas a partir de 2009. Os resultados indicam algumas características: em relação ao número de questões de física, se mantém aproximadamente constante; é difícil fazer uma correspondência unívoca entre a questão e a habilidade da Matriz de Referência (o que é exigido pela TRI); a distribuição dos objetos de conhecimento não atende a uma distribuição dos conteúdos ensinados no nível médio; os itens praticamente não envolvem a avaliação da capacidade de resolver problemas, entre outras observações apresentadas.

Finalizando, então, essa análise, pode-se concluir que a Matriz de Referência de Ciências da Natureza apresenta sérios problemas em sua construção. O exame do ENEM, baseado nesta matriz, por se constituir em uma avaliação de larga escala que é utilizada para o ingresso na maior parte das universidades públicas do Brasil, tende a médio e longo prazo, se não modificada, a promover graves distorções no currículo de Ensino de Física Brasileiro. Faz-se então necessário um estudo minucioso e detalhado que leve a construção de uma matriz que dialogue os currículos e com os professores.

CAPÍTULO 6 – CONSIDERAÇÕES FINAIS

A avaliação ocupa um papel central em qualquer processo de ensino e aprendizagem. No entanto, ela normalmente é concebida e elaborada em função da continuidade e da tradição, isto é, o professor que realiza a avaliação, na maioria das vezes, a faz através da repetição dos processos por ele vivenciados quando ainda era aluno. Logo, a existência de recursos pedagógicos que informem, orientem e promovam a reflexão desses profissionais frente a uma ferramenta tão rica e importante torna-se fundamental para a melhoria do processo educativo.

Neste contexto pode-se constatar que é possível a realização de medidas na área de ciências sociais e, ainda mais, é possível, através da operacionalização de conceitos, a mensuração de aspectos de conceitos abstratos como o de aprendizagem em física. Ainda, focando-se no estudo das avaliações em larga escala, em especial as que utilizam da Teoria de Resposta ao Item como metodologia de análise, verificou-se que algumas características destas avaliações, como a independência dos itens da prova em relação ao grupo examinado ou a não dependência do escore obtido em relação ao teste aplicado, permitem que elas sirvam de comparação entre grupos diferentes, em lugares e momentos distintos.

Neste cenário, buscando-se uma melhor compreensão destes processos avaliativos, desenvolveu-se o trabalho de pesquisa desta dissertação que se dividiu em duas etapas distintas:

Na primeira, aplicou-se um questionário que se propõe medir a aprendizagem dos conceitos de física térmica [Yeo e Zadnik 2001] a alunos do ensino médio. Da análise desses dados [Gonçalves Junior e Barroso 2011], concluiu-se que os resultados dos alunos brasileiros são similares aos relatados com estudantes australianos, inicialmente apresentando uma melhora de acordo com o nível de instrução formal recebida. No entanto, quando é feita uma análise mais detalhada das respostas dos alunos, utilizando-se para isso de estatística tradicional e da construção de curvas características desses itens, percebe-se

que as respostas dos estudantes brasileiros indicam a persistência de conceitos não científicos em suas concepções, mesmo após a instrução formal.

Procedeu-se em seguida, na continuação da análise deste questionário, à construção das Curvas Características dos Itens de suas questões, podendo-se observar que essas curvas são boas ferramentas para análise destes, mesmo quando estes itens não foram construídos para serem analisados por este método. A comparação entre as curvas características construídas para os itens possibilitou a identificação da dificuldade e do nível de discriminação destes itens e também revelou a existência de itens cuja curva característica empírica não apresenta a forma prevista habitualmente nos modelos da TRI, o que não possibilitam a análise através das ICC's, uma vez que, ao se ajustar a curva para ele, não se obteve uma curva em S, mas quase uma reta.

Apesar do foco desta primeira parte do trabalho não ser a determinação de concepções não científicas dos estudantes, foi possível constatar uma dificuldade em relação à Lei Zero da Termodinâmica, o conceito de equilíbrio térmico: grande parte dos alunos associam os processos de mudança de fase à uma temperatura (de fusão do gelo ou de ebulição da água) e não à continuidade de trocas de calor após atingir essa temperatura.

Na segunda etapa do trabalho, fez-se uma análise do Exame Nacional do Ensino Médio. A análise qualitativa elaborada destacou, primeiramente, a importância das modificações sofridas pelo exame a partir do ano de 2009, com a utilização da TRI como metodologia de análise. Esta medida possibilitou, a partir daí, a comparação entre os resultados obtidos.

Foi possível traçar um perfil desse exame, que se caracteriza por apresentar itens de caráter predominantemente qualitativo, extensos para o tempo disponível para resolvê-los e, de uma forma geral, bem contextualizados.

Durante a análise, outros dois pontos se destacaram. O primeiro deles é o fato de que os itens podem representar mais de uma habilidade.

O segundo é o fato de que a distribuição não homogênea dos itens nos objetos específicos de conteúdo, privilegiando os temas que abordam conceitos de eletricidade e termodinâmica. Este fator, a médio e longo prazo, pode produzir sérias distorções no currículo de física do ensino médio no Brasil.

A análise desta segunda etapa do trabalho apontou para a necessidade de reformulação da Matriz de Referência do ENEM, para que os objetos de conhecimento estejam mais alinhados com os apresentados no ensino médio, para que a distribuição de habilidades e competências se dê de forma mais homogênea, permitindo que o exame se torne um instrumento de maior utilidade para compor indicadores da qualidade da educação no ensino médio do país.

E finalmente, já que as avaliações governamentais de larga escala colocam muitas vezes em cheque o professor em relação às suas práticas avaliativas, compreender, conhecer e refletir sobre as características destas avaliações torna-se primordial para o aprimoramento profissional deste professor.

Foi com esta ideia em mente, então, que foi proposto um texto para professores de ensino médio que faça uma discussão dessas características de forma a permitir que o professor continue se desenvolvendo profissionalmente e melhorando a qualidade do processo de ensino aprendizagem.

Referências Bibliográficas

[Arons 1997] A. B. ARONS, The Concept of Heat. In Teaching Introductory Physics, Contents of Part III. Washington: John Wiley & Sons, 1997, p. 67-82.

[Allen e Yen 2002] Mary J. Allen, Wendy M. Yen, Introduction to Measurement Theory, Long Grove: Waveland Press, 2002, p. 2.

[Alonso e Finn 1972] M. Alonso, E. J. Finn, Física: Um Curso Universitário, Tradução (coord.) Giorgio Moscati. São Paulo: Ed. Edgard Blücher LTDA, 1972, p. 13.

[Babbie 2005] Earl Babbie, Métodos de pesquisas de Survey. Tradução de Guilherme Cezarino, Belo Horizonte: UFMG, 1999, p. 57-75 e p.179-184.

[Barroso e Borgo 2010] Marta Feijó Barroso e Igor Borgo. Jornada no Sistema Solar, Revista Brasileira de Ensino de Física, v. 32, n. 2, 2052, 2010. Disponível em <http://www.sbfisica.org.br/rbef/pdf/322502.pdf>. Acesso em 25/10/2011.

[Barroso e Franco 2008] Marta Feijó Barroso e Creso Franco. Avaliações Educacionais: O PISA e o Ensino de Ciências. XI Encontro de Pesquisa em Ensino de Física, Curitiba, 2008.

[Black 1998] Paul Black, Testing: Friend or Foe? The theory and practice of assessment and testing. London: RoutledgeFalmer, 1998, p. 24-36.

[Bonamino, Coscarelli e Franco 2002] A. Bonamino, C. Coscarelli, C. Franco, Avaliação e Letramento: Concepções de Aluno Letrado Subjacentes ao SAEB e ao PISA, Educação e Sociedade, Campinas, vol. 23, n. 81, p. 91-113, dezembro 2002. Disponível em: <http://www.cedes.unicamp.br>. Acesso em: 25 junho 2011.

[Britton e Schneider 2007] E.D. Britton e S.A. Schneider, Large-Scale Assessments in Science Education. In: S. K. Abell e N.G. Lederman (ed.), Handbook of Research on Science Education. Mahwah: Laurence Erlbaum Associates, 2007, p. 1007-1040.

[Clough e Driver 1985] E. E. Clough, R. Driver, Secondary student's conceptions of the Conduction of heat: bringing together scientific and personal views. Physycs Education, v.20, n.4, p. 176-182, julho, 1985.

[Duit 2009] R. Duit, Bibliography – STCSE - Students' and Teachers' Conceptions and Science Education. Compilada por Reinders Duit. Disponível em: <<http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html>> consultado em 27/03/2009.

[Gonçalves Junior e Barroso 2011a] W. P. Gonçalves Junior, M. F. Barroso, O Que Alunos Entendem a Respeito de Conceitos de Básicos de Física Térmica, XIX Simpósio Nacional de Ensino de Física, Manaus, 2011.

[Gonçalves Junior e Barroso 2011b] W. P. Gonçalves Junior, M. F. Barroso, Conceitos de Física Térmica: Estudo por Meio de um Questionário, XIII Encontro de Pesquisa em Ensino de Física, Foz do Iguaçu, 2011.

[Harrison, Grayson e Treagust 1999] A. G. Harrison, J. D. Grayson, D. F. Treagust, Investigating a Grade 11 Student's Evolving Conceptions of Heat and Temperature. Journal of Research in Science Teaching, v. 36, n.1, p. 55-87, janeiro, 1999.

[Hambleton 1991] R. Hambleton, H. Swaminathan, H. J. Rogers, Fundamentals of Item Response Theory, Newbury Park: SAGE Publications, 1991, p. 1-31.

[Hambleton 1993] R. K Hambleton, Principles and Selected Applications of Item Response Theory. In: R.L. Linn, Educational Measurement, Third Edition.

American Council on Education Series on Higher Education, Phoenix: Orix Press, 1993. P. 147-200.

[Hestenes 1992] D. Hestenes , M. Wells, G. Swackhamer, Force concept inventory. The Physics Teacher 30: 141-166.

[Keeves 1992] J. P.Keeves, The IEA Study of Science III: Changes in Science Education and Achievement 1970 to 1984, England: Pergamon, 11992.

[Klein e Fontanive 1995] Ruben Klein Nilma e Santos Fontanive, Avaliação em Larga Escala: Uma Proposta Inovadora, Em Aberto, Brasília, ano 15, n. 66, p. 28-34, 1995. Disponível em: <<http://emaberto.inep.gov.br/index.php/emaberto/article/viewFile/995/899>>. Acesso: 25/06/2011.

[Machado 2010] Ledo Vaccaro Machado, Avaliações de Larga Escala e Proficiência Matemática. Dissertação apresentada ao Programa de Pós-Graduação em Ensino de Matemática. Disponível em <http://www.pg.im.ufrj.br/pemat/26%20Ledo%20Vaccaro.pdf>. Consultada em 30/09/2011.

[Nentwig 2009] P. Nentwig, S. Roennebeck, K. Schoeps, S. Rumann, C. Carstensen, Performance and Levels of Contextualization in a Selection of OECD Countries in Pisa 2006, Journal of Research in Science Teaching, vol. 46, n. 8, p. 897-908, 2009.

[OECD 2000] Organisation for Economic Co-Operation and Development. Measuring Students Knowledge and Skills: the Pisa 2000 Assessment of Reading, Mathematical and Scientific Literacy. Paris : OECD, 2000.

[OECD 2009] Organisation for Economic Co-Operation and Development, PISA Data Analysis Manual -SPSS® SECOND EDITION, Paris : OECD, 2009. Disponível em: <http://www.oecd.org/document/19/0,3746,en_2649_3584562_1_48577747_1_1_1_1,00.html>. Consultado em 13/11/2011.

[Pasquali 1996] L. Pasquali, Teoria e Métodos de Medida em Ciências do Comportamento, Brasília: INEP, 1996, p. 71-72.

[Pasquali 2003] L. Pasquali, Fundamentos da Teoria da Resposta ao Item – TRI, v.2 n.2, Brasília, 2003, p. 99-110. Disponível em: <<http://pepsic.bvsalud.org/pdf/avp/v2n2/v2n2a02.pdf>>. Acesso: 12/07/2011.

[Pasquali 2009] L. Pasquali, Psicometria: Teoria dos Testes na Psicologia e na Educação, Brasília: Ed. Vozes, 2009, p. 52-90.

[PCN+ 2002] Ministério da Educação (MEC), Secretaria de Educação Média e Tecnológica (Semtec). PCN + Ensino médio: orientações educacionais complementares aos Parâmetros Curriculares Nacionais – Ciências da Natureza, Matemática e suas Tecnologias. Brasília: MEC/Semtec, 2002. p. 45-62.

[Perrenoud 1999] Philippe Perrenoud, Avaliação: da excelência à regulação das aprendizagens – entre duas lógicas, trad. Patrícia Chittoni Ramos, Porto Alegre: Artes Médicas Sul, 1999, p. 10-17.

[Portal INEP 2011] INEP. Disponível em: <www.inep.gov.br>. Acesso: 29/06/2011.

[Viennot 1998] L. VIENNOT, Experimental Facts and Ways of Reasoning in Thermodynamics: Learners' Common Approach, in: TIBERGHIE, Andrée; JOSSEM, E. Leonard; BAROJAS, Jorge: Connecting Research in Physics Education with Teacher Education, International Commission on Physics Education, 1998. Disponível em: <<http://www.physics.ohio-state.edu/~jossem/ICPE/C3.html>> Acesso em: 27/11/2010.

[Wagemaker, 2008] H. Wagemaker, IEA's International Assessments in Reading, Mathematics and Science: Specifications, Benefits and Resource Requirements. Otago, 2008. Disponível em: <<http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=1826618>>. Acesso: 06/09/2011.

[Yeo e Zadnik 2001] S. YEO, M. ZADNIK, Introductory Thermal Concept Evaluation: Assessing Student's Understanding, The Physics Teacher, v. 39, p. 496-504, novembro, 2001.

[Young 1962] H. D. YOUNG, Statistical Treatment of Experimental Data, New York: McGraw-Hill Book Company, 1962, p. 39-126.

Anexo 1

Documento Básico

Introdução

O presente documento descreve o Exame Nacional do Ensino Médio – ENEM, instituído pelo Instituto Nacional de Estudos e Pesquisas Educacionais — INEP, em 1998, para ser aplicado aos alunos concluintes e aos egressos deste nível de ensino. O ENEM será realizado anualmente, com o objetivo fundamental de avaliar o desempenho do aluno ao término da escolaridade básica, para aferir o desenvolvimento de competências fundamentais ao exercício pleno da cidadania.

As tendências internacionais, tanto em realidades mais próximas da nossa como nas mais distantes, acentuam a importância da formação geral na educação básica, não só para a continuidade da vida acadêmica como, também, para uma atuação autônoma do sujeito na vida social, com destaque à sua inserção no mercado de trabalho, que se torna mais e mais competitivo. Esta formação deve ser compreendida como uma sólida aquisição dos conteúdos tradicionais das ciências e das artes associada ao desenvolvimento de estruturas capazes de operacionalizá-los no enfrentamento de problemas apresentados pela realidade social, cada vez mais complexa, e numa dinâmica de tempo progressivamente acelerada.

Esta rapidez com que as mudanças sociais se processam e alteram nossa vida cotidiana impõe um padrão mais elevado para a escolaridade básica, e o projeto pedagógico da escola deve objetivar o desenvolvimento de competências com as quais os alunos possam assimilar informações e utilizá-las em contextos adequados, interpretando códigos e linguagens e servindo-se dos conhecimentos adquiridos para a tomada de decisões autônomas e socialmente relevantes.

Estas premissas já estão delineadas na atual Lei de Diretrizes e Bases da Educação Nacional – LDB, que introduz profundas transformações no ensino médio, desvinculando-o do vestibular, ao flexibilizar os mecanismos de acesso ao ensino superior, e, principalmente, delineando o perfil de saída do aluno da escolaridade básica, ao estipular que o educando, ao final do ensino médio, demonstre:

- “I – domínio dos princípios científicos e tecnológicos que presidem a produção moderna;
- II – conhecimento das formas contemporâneas de linguagem;
- III – domínio dos conhecimentos de Filosofia e de Sociologia necessários ao exercício da cidadania.”

No âmbito dessas mudanças, a LDB determina, inclusive, que a União organize processo nacional de avaliação do rendimento escolar, para todos os níveis de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino (art. 9º, VI).

É nessa perspectiva que o INEP vem realizando o ENEM, para o universo de alunos concluintes e de egressos deste nível de ensino.

Esse exame difere de outras avaliações já propostas pelo Ministério da Educação. Centra-se na avaliação de desempenho por competências e vincula-se a um conceito mais abrangente e estrutural da inteligência humana. O exame é constituído de uma prova única

e abrange as várias áreas de conhecimento em que se organizam as atividades pedagógicas da escolaridade básica no Brasil.

Para estruturar o exame, concebeu-se uma matriz com a indicação de competências e habilidades associadas aos conteúdos do ensino fundamental e médio que são próprias ao sujeito na fase de desenvolvimento cognitivo, correspondente ao término da escolaridade básica. Tem como referência a LDB, os Parâmetros Curriculares Nacionais (PCN), a Reforma do Ensino Médio, bem como os textos que sustentam sua organização curricular em Áreas de Conhecimento, e, ainda, as Matrizes Curriculares de Referência para o SAEB.

A realização anual do ENEM junto aos alunos que completaram ou estão completando a escolaridade básica poderá fornecer uma imagem realista e sempre atualizada da educação no Brasil.

1. Objetivos

O ENEM será realizado anualmente, com o objetivo fundamental de avaliar o desempenho do aluno ao término da escolaridade básica, para aferir o desenvolvimento de competências fundamentais ao exercício pleno da cidadania. Pretende, ainda, alcançar os seguintes objetivos específicos:

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

2. Características do Exame

2.1. Inscrição

As inscrições serão realizadas em datas e locais a serem definidos, anualmente, em portaria do INEP.

As inscrições serão feitas por instituição contratada para este fim, devendo o interessado preencher corretamente a Ficha de Inscrição, responsabilizando-se pelas informações nelas fornecidas, e efetuar o pagamento da taxa de inscrição.

Para efetuar a inscrição, o interessado deverá portar documento de identidade e cópia do mesmo, para anexá-la à sua ficha de inscrição

Poderão ser aceitas inscrições coletivas patrocinadas por instituições públicas ou privadas, sendo vedado o patrocínio de empresas relacionadas à produção e comercialização de bebidas alcoólicas e/ou fumo. No caso de inscrição patrocinada, a instituição interessada:

- I. solicitar à instituição contratada para realizar a inscrição formulários de inscrição ao exame em quantidade correspondente ao número de candidatos que desejar patrocinar; e
- II. de posse dos formulários de inscrição, providenciar e responsabilizar-se pelo preenchimento dos mesmos por seus patrocinados e pelo pagamento do valor correspondente às taxas de inscrição dos candidatos que estiver patrocinando, junto à instituição contratada.

O portador de necessidades especiais, interessado em participar do ENEM, no ato da inscrição deverá, obrigatoriamente, declarar o tipo de necessidade a que se refere, sob pena de não ter atendimento apropriado.

Para o portador de necessidade especial amblópe ou cego, a instituição responsável pela aplicação do exame deverá preparar prova e atendimento especial. Aos amblópes, serão oferecidas provas ampliadas com tamanho de letra correspondente ao corpo 24, e, aos cegos serão disponibilizadas provas em braile.

Aos portadores de deficiência física com séria dificuldade de locomoção, serão oferecidas salas de fácil acesso, e aos participantes incapazes de efetuar a marcação do cartão-resposta, será oferecido auxílio para transcrição.

2.2. Prova

O exame será constituído por uma prova única contendo 63 (sessenta e três) questões objetivas de múltipla escolha e uma proposta para redação. As questões objetivas e a redação destinam-se a avaliar as competências e habilidades desenvolvidas pelos participantes ao longo da escolaridade básica, a partir de uma Matriz de Competências especialmente desenvolvida para estruturar o exame.

2.3. Data de Realização do Exame

O exame será realizado no último domingo de agosto, com início às 13 (treze) horas (horário de Brasília) e terá duração de 5 (cinco) horas.

Considerando o horário de Brasília-DF para todo o território Nacional, os portões de acesso aos locais de prova serão abertos às 12 (doze) horas e fechados às 13 (treze) horas impreterivelmente, não sendo permitida a entrada do inscrito que se apresentar após o horário estipulado.

2.4. Locais de Realização do Exame

O exame será realizado anualmente, com aplicação descentralizada das provas, nas capitais, no Distrito Federal e nos demais municípios que, a critério do MEC/INEP, ofereçam condições estratégicas para sua realização de modo a atender adequadamente às demandas do ensino médio e de seus egressos e dentro da estrutura de operacionalização possível, definida para o ENEM, em portaria.

3. Participantes

O ENEM tem caráter voluntário e dele podem participar, mediante inscrição, os concluintes do ensino médio, no ano de realização do exame, e também os que já o concluíram em anos anteriores, em qualquer de suas modalidades. É direito do participante realizar o ENEM quantas vezes for de seu interesse.

4. Manual do Inscrito

O Manual do Inscrito será remetido ao endereço indicado na ficha de inscrição, juntamente com a confirmação da inscrição e conterá as informações gerais sobre o Exame, as competências e habilidades a serem avaliadas, os critérios de avaliação de desempenho do participante nas duas partes da prova, bem como o questionário socioeconômico com sua respectiva folha de respostas, apropriada para leitura ótica.

O inscrito no ENEM deverá responder ao questionário socioeconômico, parte integrante do Manual do Inscrito, que permitirá traçar o perfil dos participantes e também desenvolver estudos contextuais sobre os resultados anuais do Exame.

Todos os quesitos do Questionário Socioeconômico deverão ser respondidos na folha de resposta para este fim, que deverá ser entregue no dia e local de realização da prova.

5. Matriz de Competências

A Matriz de Competências foi desenvolvida para estruturar o ENEM, a fim de definir claramente seus pressupostos e delinear suas características operacionais.

A Matriz foi construída por um grupo de profissionais da educação – especialistas em psicologia do desenvolvimento, pesquisadores e professores das diferentes áreas de conhecimento e especialistas em psicometria –, a partir de um projeto elaborado e coordenado pelo INEP.

A concepção de conhecimento subjacente a essa matriz pressupõe colaboração, complementaridade e integração entre os conteúdos das diversas áreas do conhecimento presentes nas propostas curriculares das escolas brasileiras de ensino fundamental e médio e considera que conhecer é construir e reconstruir significados continuamente, mediante o estabelecimento de relações de múltipla natureza, individuais e sociais.

O modelo da Matriz contempla a indicação das competências e habilidades gerais próprias do aluno, na fase de desenvolvimento cognitivo correspondente ao término da escolaridade básica, associadas aos conteúdos do ensino fundamental e médio, e considera, como referências norteadoras, o texto da Lei de Diretrizes e Bases da Educação Nacional (LDB), os Parâmetros Curriculares Nacionais (PCN), os textos da Reforma do Ensino Médio e as Matrizes Curriculares de Referência para o SAEB.

Competências são as modalidades estruturais da inteligência, ou melhor, ações e operações que utilizamos para estabelecer relações com e entre objetos, situações, fenômenos e pessoas que desejamos conhecer. As *habilidades* decorrem das competências adquiridas e referem-se ao plano imediato do “saber fazer”. Por meio das ações e operações, as habilidades aperfeiçoam-se e articulam-se, possibilitando nova reorganização das competências.

A Matriz pressupõe, ainda, que a competência de ler, compreender, interpretar e produzir textos, no sentido amplo do termo, não se desenvolve unicamente na aprendizagem da Língua Portuguesa, mas em todas as áreas e disciplinas que estruturam as atividades pedagógicas na escola. O aluno deve, portanto, demonstrar, concomitantemente, possuir instrumental de comunicação e expressão adequado tanto para a compreensão de um problema matemático quanto para a descrição de um processo físico, químico ou biológico e, mesmo, para a percepção das transformações de espaço/tempo da história, da geografia e da literatura.

A partir das competências cognitivas globais, identificou-se o elenco de habilidades correspondentes, e a matriz assim construída fornece indicações do que se pretende valorizar nessa avaliação, servindo de orientação para a elaboração de questões que envolvam as diferentes áreas do conhecimento.

Busca-se, dessa maneira, verificar como o conhecimento assim construído pode ser efetivado pelo participante por meio da demonstração de sua autonomia de julgamento e de ação, de atitudes, valores e procedimentos diante de situações-problema que se aproximem o máximo possível das condições reais de convívio social e de trabalho individual e coletivo.

Todas as situações de avaliação estruturam-se de modo a verificar se o participante é capaz de ler e interpretar textos de linguagem verbal, visual (fotos, mapas, pinturas, gráficos, entre outros) e enunciados:

- identificando e selecionando informações centrais e periféricas;
- inferindo informações, temas, assuntos, contextos;
- justificando a adequação da interpretação;
- compreendendo os elementos implícitos de construção do texto, como organização, estrutura, intencionalidade, assunto e tema;
- analisando os elementos constitutivos dos textos, de acordo com sua natureza, organização ou tipo;
- comparando os códigos e linguagens entre si, reelaborando, transformando e reescrevendo (resumos, paráfrases e relatos).

ENEM - COMPETÊNCIAS

- I. Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica.
- II. 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. Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.
- IV. Relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.
- V. Recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.

ENEM – HABILIDADES

1. Dada a descrição discursiva ou por ilustração de um experimento ou fenômeno, de natureza científica, tecnológica ou social, identificar variáveis relevantes e selecionar os instrumentos necessários para realização ou interpretação do mesmo.
2. Em um gráfico cartesiano de variável socioeconômica ou técnico-científica, identificar e analisar valores das variáveis, intervalos de crescimento ou decréscimo e taxas de variação.
3. Dada uma distribuição estatística de variável social, econômica, física, química ou biológica, traduzir e interpretar as informações disponíveis, ou reorganizá-las, objetivando interpolações ou extrapolações.
4. Dada uma situação-problema, apresentada em uma linguagem de determinada área de conhecimento, relacioná-la com sua formulação em outras linguagens ou vice-versa.
5. A partir da leitura de textos literários consagrados e de informações sobre concepções artísticas, estabelecer relações entre eles e seu contexto histórico, social, político ou cultural, inferindo as escolhas dos temas, gêneros discursivos e recursos expressivos dos autores.
6. Com base em um texto, analisar as funções da linguagem, identificar marcas de variantes lingüísticas de natureza sociocultural, regional, de registro ou de estilo, e explorar as relações entre as linguagens coloquial e formal.

7. Identificar e caracterizar a conservação e as transformações de energia em diferentes processos de sua geração e uso social, e comparar diferentes recursos e opções energéticas.
8. Analisar criticamente, de forma qualitativa ou quantitativa, as implicações ambientais, sociais e econômicas dos processos de utilização dos recursos naturais, materiais ou energéticos.
9. Compreender o significado e a importância da água e de seu ciclo para a manutenção da vida, em sua relação com condições socioambientais, sabendo quantificar variações de temperatura e mudanças de fase em processos naturais e de intervenção humana.
10. Utilizar e interpretar diferentes escalas de tempo para situar e descrever transformações na atmosfera, biosfera, hidrosfera e litosfera, origem e evolução da vida, variações populacionais e modificações no espaço geográfico.
11. Diante da diversidade da vida, analisar, do ponto de vista biológico, físico ou químico, padrões comuns nas estruturas e nos processos que garantem a continuidade e a evolução dos seres vivos.
12. Analisar fatores socioeconômicos e ambientais associados ao desenvolvimento, às condições de vida e saúde de populações humanas, por meio da interpretação de diferentes indicadores.
13. Compreender o caráter sistêmico do planeta e reconhecer a importância da biodiversidade para preservação da vida, relacionando condições do meio e intervenção humana.
14. Diante da diversidade de formas geométricas planas e espaciais, presentes na natureza ou imaginadas, caracterizá-las por meio de propriedades, relacionar seus elementos, calcular comprimentos, áreas ou volumes, e utilizar o conhecimento geométrico para leitura, compreensão e ação sobre a realidade.
15. Reconhecer o caráter aleatório de fenômenos naturais ou não e utilizar em situações-problema processos de contagem, representação de frequências relativas, construção de espaços amostrais, distribuição e cálculo de probabilidades.
16. Analisar, de forma qualitativa ou quantitativa, situações-problema referentes a perturbações ambientais, identificando fonte, transporte e destino dos poluentes, reconhecendo suas transformações; prever efeitos nos ecossistemas e no sistema produtivo e propor formas de intervenção para reduzir e controlar os efeitos da poluição ambiental.
17. Na obtenção e produção de materiais e de insumos energéticos, identificar etapas, calcular rendimentos, taxas e índices, e analisar implicações sociais, econômicas e ambientais.

18. Valorizar a diversidade dos patrimônios etnoculturais e artísticos, identificando-a em suas manifestações e representações em diferentes sociedades, épocas e lugares.
19. Confrontar interpretações diversas de situações ou fatos de natureza histórico-geográfica, técnico-científica, artístico-cultural ou do cotidiano, comparando diferentes pontos de vista, identificando os pressupostos de cada interpretação e analisando a validade dos argumentos utilizados.
20. Comparar processos de formação socioeconômica, relacionando-os com seu contexto histórico e geográfico.
21. Dado um conjunto de informações sobre uma realidade histórico-geográfica, contextualizar e ordenar os eventos registrados, compreendendo a importância dos fatores sociais, econômicos, políticos ou culturais.

6. Análise de Desempenho

O desempenho do participante será avaliado nas duas partes da prova (objetiva e redação), valendo 100 pontos cada uma delas. Esse desempenho será qualificado de acordo com as premissas teóricas da Matriz de Competências que se refere às possibilidades totais da cognição humana na fase de desenvolvimento próprio aos participantes do ENEM – jovens e adultos. Essa qualificação será expressa nas seguintes faixas de desempenho: insuficiente a regular, que corresponde às notas entre 0 a 40, inclusive; regular a bom, que corresponde às notas entre 40 a 70, inclusive; e de bom a excelente, que corresponde às notas entre 70 a 100.

6.1. Modelo de Análise de Desempenho na Parte Objetiva da Prova

A parte objetiva da prova será constituída de 63 (sessenta e três) questões de múltipla escolha de igual valor, avaliada numa escala de zero a 100 (cem) pontos, e gera uma nota global que corresponde à soma dos pontos atribuídos às questões acertadas.

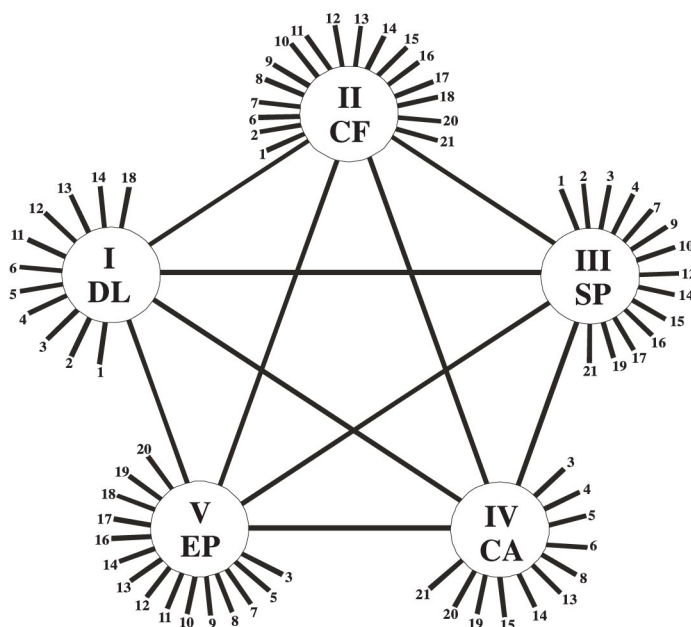
As cinco competências que são avaliadas no ENEM na parte objetiva da prova expressam-se por meio de 21 habilidades. Cada uma das 21 habilidades será medida três vezes (três questões para cada habilidade).

A interpretação dessa nota global será estruturada a partir de cada uma das cinco competências, pelas relações estabelecidas com as respectivas habilidades e as questões a elas relacionadas, gerando também para cada competência, uma nota de 0 a 100, conforme modelo a seguir:

Competências:

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

Habilidades: 1 a 21



6.2. Modelo de análise de desempenho na Redação

A redação deverá ser estruturada na forma de texto em prosa do tipo dissertativo-argumentativo, a partir da proposta de um tema de ordem social, científica, cultural ou política.

Na redação, também serão avaliadas as cinco competências da Matriz do ENEM, referidas à produção de um texto. Cada uma das competências será avaliada numa escala de 0 a 100 pontos.

Caso o participante não desenvolva o tema e a estrutura solicitados, será atribuída a nota ZERO à competência II da redação, o que anula a correção das demais competências da redação. A nota global da redação, neste caso, será ZERO.

A nota global da redação será dada pela média aritmética das notas atribuídas a cada uma das cinco competências específicas da redação.

As cinco competências avaliadas na redação são as mesmas avaliadas na parte objetiva da prova, traduzidas para uma situação específica de produção de texto, conforme especificado a seguir.

Competências do ENEM

Parte Objetiva	Redação
I. Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica.	I. Demonstrar domínio da norma culta da língua escrita.
II. 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.	II. Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo.
III. Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.	III. Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista.
IV. Relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.	IV. Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação.
V. Recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.	V. Elaborar proposta de solução para o problema abordado, mostrando respeito aos valores humanos e considerando a diversidade sociocultural.

Na **competência I**, espera-se que o participante escolha o registro adequado a uma situação formal de produção de texto escrito. Na avaliação, serão considerados os fundamentos gramaticais do texto escrito, refletidos na utilização da norma culta em aspectos como: sintaxe de concordância, regência e colocação; pontuação; flexão; ortografia; e adequação de registro demonstrada, no desempenho linguístico, de acordo com a situação formal de produção exigida.

O eixo da **competência II** reside na compreensão do tema que instaura uma problemática a respeito da qual se pede um texto escrito em prosa do tipo dissertativo-argumentativo em prosa. Por meio deste tipo de texto, analisam-se, interpretam-se e

relacionam-se dados, informações e conceitos amplos, tendo-se em vista a construção de uma argumentação, em defesa de um ponto de vista.

Na **competência III**, procura-se avaliar como o participante, em uma situação formal de interlocução, seleciona, organiza, relaciona e interpreta os dados, informações e conceitos necessários para defender sua perspectiva sobre o tema proposto.

Na **competência IV**, avalia-se a utilização de recursos coesivos da modalidade escrita, com vistas à adequada articulação dos argumentos, fatos e opiniões selecionados para a defesa de um ponto de vista sobre o tema proposto. Serão considerados os mecanismos lingüísticos responsáveis pela construção da argumentação na superfície textual, tais como: coesão referencial; coesão lexical (sinônimos, hiperônimos, repetição, reiteração); e coesão gramatical (uso de conectivos, tempos verbais, pontuação, seqüência temporal, relações anafóricas, conectores intervocabulares, intersetenciais, interparágrafos).

Na **competência V**, verifica-se como o participante indicará as possíveis variáveis para solucionar a problemática desenvolvida, quais propostas de intervenção apresentou, qual a relação destas com o projeto desenvolvido sobre o tema proposto e a qualidade destas propostas, mais genéricas ou específicas, tendo por base a solidariedade humana e o respeito à diversidade de pontos de vista, eixos de uma sociedade democrática.

7. Resultados

7.1. Resultados Individuais

Os participantes do Exame receberão, no endereço indicado na Ficha de Inscrição, o Boletim Individual de Resultados.

No Boletim Individual de Resultados constarão duas notas, uma para a parte objetiva e outra para a redação e, ainda, uma interpretação dos resultados obtidos para cada uma das cinco competências avaliadas nas duas partes da prova, de acordo com o modelo estabelecido na Matriz de Competência do ENEM.

Os resultados individuais do ENEM não serão divulgados por meio de publicação ou instrumentos similares, podendo, todavia, as instituições neles interessadas—estabelecimentos de ensino pós-médio e superior e instituições do mercado de trabalho— a eles ter acesso desde que obtenham a autorização do participante.

O participante poderá autorizar a utilização de seus resultados obtidos no ENEM a todos os segmentos sociais que a ele interessar. O MEC/INEP deverá confirmar oficialmente estes resultados, quando acionado formalmente, desde que obtenham autorização formal do participante.

7.2. Resultados para as Instituições de Ensino Pós-Médio e Superior

As instituições de ensino pós-médio e superior que desejarem utilizar os resultados individuais do ENEM, como critério de seleção às suas vagas, deverão encaminhar formalmente ao MEC/INEP a sua solicitação.

Os participantes deverão fornecer o número de inscrição do ENEM às instituições, o que caracterizará sua autorização para uso de seus resultados.

As instituições de ensino pós-médio ou superior que desejarem utilizar os resultados individuais do ENEM deverão planejar a inscrição de seu processo seletivo de modo a atender às exigências cadastrais dos sistemas coletores de dados do MEC/INEP, principalmente no que se refere ao número de inscrição do ENEM (12 dígitos), pois sem o mesmo não será possível fornecer resultados.

As instituições que não dispuserem do número de inscrição dos participantes não receberão os resultados individuais dos mesmos.

7.3. Resultados para as Instituições de Ensino Médio

Resguardado o sigilo dos resultados individuais, o INEP poderá elaborar o Boletim de Resultados da Escola, com a análise de desempenho global do conjunto de concluintes do ensino médio da instituição de ensino interessada, desde que:

- I. encaminhem ao MEC/INEP solicitação formal;
- II. declarem formalmente que pelo menos 90% (noventa por cento) de seus alunos tenham participado do ENEM;
- III. comprovem o recolhimento, em favor do INEP, no caso de instituições privadas, da importância de R\$5,00 (cinco reais) por aluno. As instituições públicas estarão isentas do recolhimento dessa importância.

As instituições de ensino médio que não dispuserem do número de inscrição dos participantes não receberão o Boletim de Resultados da Escola, ainda que atendam às especificações acima.

O MEC/INEP não disponibilizará os resultados individuais dos participantes para as instituições de ensino médio.

Maiores Informações

Ligue 0800-616161

Consulte o site <http://www.inep.gov.br/enem>

Anexo 2



MINISTÉRIO DA EDUCAÇÃO
INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA

MATRIZ DE REFERÊNCIA PARA O ENEM 2009

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.

Matriz de Referência de Linguagens, Códigos e suas Tecnologias

Competência de área 1 - Aplicar as tecnologias da comunicação e da informação na escola, no trabalho e em outros contextos relevantes para sua vida.

H1 - Identificar as diferentes linguagens e seus recursos expressivos como elementos de caracterização dos sistemas de comunicação.

H2 - Recorrer aos conhecimentos sobre as linguagens dos sistemas de comunicação e informação para resolver problemas sociais.

H3 - Relacionar informações geradas nos sistemas de comunicação e informação, considerando a função social desses sistemas.

H4 - Reconhecer posições críticas aos usos sociais que são feitos das linguagens e dos sistemas de comunicação e informação.

Competência de área 2 - Conhecer e usar língua(s) estrangeira(s) moderna(s) como instrumento de acesso a informações e a outras culturas e grupos sociais*.

H5 – Associar vocábulos e expressões de um texto em LEM ao seu tema.

H6 - Utilizar os conhecimentos da LEM e de seus mecanismos como meio de ampliar as possibilidades de acesso a informações, tecnologias e culturas.

H7 – Relacionar um texto em LEM, as estruturas linguísticas, sua função e seu uso social.

H8 - Reconhecer a importância da produção cultural em LEM como representação da diversidade cultural e linguística.

***A área 2 será incluída apenas a partir de 2010**

Competência de área 3 - Compreender e usar a linguagem corporal como relevante para a própria vida, integradora social e formadora da identidade.

H9 - Reconhecer as manifestações corporais de movimento como originárias de necessidades cotidianas de um grupo social.

H10 - Reconhecer a necessidade de transformação de hábitos corporais em função das necessidades cinestésicas.

H11 - Reconhecer a linguagem corporal como meio de interação social, considerando os limites de desempenho e as alternativas de adaptação para diferentes indivíduos.

Competência de área 4 - Compreender a arte como saber cultural e estético gerador de significação e integrador da organização do mundo e da própria identidade.

H12 - Reconhecer diferentes funções da arte, do trabalho da produção dos artistas em seus meios culturais.

H13 - Analisar as diversas produções artísticas como meio de explicar diferentes culturas, padrões de beleza e preconceitos.

H14 - Reconhecer o valor da diversidade artística e das inter-relações de elementos que se apresentam nas manifestações de vários grupos sociais e étnicos.

Competência de área 5 - Analisar, interpretar e aplicar recursos expressivos das linguagens, relacionando textos com seus contextos, mediante a natureza, função, organização, estrutura das manifestações, de acordo com as condições de produção e recepção.

H15 - Estabelecer relações entre o texto literário e o momento de sua produção, situando aspectos do contexto histórico, social e político.

H16 - Relacionar informações sobre concepções artísticas e procedimentos de construção do texto literário.

H17 - Reconhecer a presença de valores sociais e humanos atualizáveis e permanentes no patrimônio literário nacional.

Competência de área 6 - Compreender e usar os sistemas simbólicos das diferentes linguagens como meios de organização cognitiva da realidade pela constituição de significados, expressão, comunicação e informação.

H18 - Identificar os elementos que concorrem para a progressão temática e para a organização e estruturação de textos de diferentes gêneros e tipos.

H19 - Analisar a função da linguagem predominante nos textos em situações específicas de interlocução.

H20 - Reconhecer a importância do patrimônio linguístico para a preservação da memória e da identidade nacional.

Competência de área 7 - Confrontar opiniões e pontos de vista sobre as diferentes linguagens e suas manifestações específicas.

H21 - Reconhecer em textos de diferentes gêneros, recursos verbais e não-verbais utilizados com a finalidade de criar e mudar comportamentos e hábitos.

H22 - Relacionar, em diferentes textos, opiniões, temas, assuntos e recursos linguísticos.

H23 - Inferir em um texto quais são os objetivos de seu produtor e quem é seu público alvo, pela análise dos procedimentos argumentativos utilizados.

H24 - Reconhecer no texto estratégias argumentativas empregadas para o convencimento do público, tais como a intimidação, sedução, comoção, chantagem, entre outras.

Competência de área 8 - Compreender e usar a língua portuguesa como língua materna, geradora de significação e integradora da organização do mundo e da própria identidade.

H25 - Identificar, em textos de diferentes gêneros, as marcas linguísticas que singularizam as variedades linguísticas sociais, regionais e de registro.

H26 - Relacionar as variedades linguísticas a situações específicas de uso social.

H27 - Reconhecer os usos da norma padrão da língua portuguesa nas diferentes situações de comunicação.

Competência de área 9 - Entender os princípios, a natureza, a função e o impacto das tecnologias da comunicação e da informação na sua vida pessoal e social, no desenvolvimento do conhecimento, associando-o aos conhecimentos científicos, às linguagens que lhes dão suporte, às demais tecnologias, aos processos de produção e aos problemas que se propõem solucionar.

H28 - Reconhecer a função e o impacto social das diferentes tecnologias da comunicação e informação.

H29 - Identificar pela análise de suas linguagens, as tecnologias da comunicação e informação.

H30 - Relacionar as tecnologias de comunicação e informação ao desenvolvimento das sociedades e ao conhecimento que elas produzem.

Matriz de Referência de Matemática e suas Tecnologias

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.

Matriz de Referência de Ciências da Natureza e suas Tecnologias

Competência de área 1 – Compreender as ciências naturais e as tecnologias a elas associadas como construções humanas, percebendo seus papéis nos processos de produção e no desenvolvimento econômico e social da humanidade.

H1 – Reconhecer características ou propriedades de fenômenos ondulatórios ou oscilatórios, relacionando-os a seus usos em diferentes contextos.

H2 – Associar a solução de problemas de comunicação, transporte, saúde ou outro, com o correspondente desenvolvimento científico e tecnológico.

H3 – Confrontar interpretações científicas com interpretações baseadas no senso comum, ao longo do tempo ou em diferentes culturas.

H4 – Avaliar propostas de intervenção no ambiente, considerando a qualidade da vida humana ou medidas de conservação, recuperação ou utilização sustentável da biodiversidade.

Competência de área 2 – Identificar a presença e aplicar as tecnologias associadas às ciências naturais em diferentes contextos.

H5 – Dimensionar circuitos ou dispositivos elétricos de uso cotidiano.

H6 – Relacionar informações para compreender manuais de instalação ou utilização de aparelhos, ou sistemas tecnológicos de uso comum.

H7 – Selecionar testes de controle, parâmetros ou critérios para a comparação de materiais e produtos, tendo em vista a defesa do consumidor, a saúde do trabalhador ou a qualidade de vida.

Competência de área 3 – Associar intervenções que resultam em degradação ou conservação ambiental a processos produtivos e sociais e a instrumentos ou ações científico-tecnológicos.

H8 – Identificar etapas em processos de obtenção, transformação, utilização ou reciclagem de recursos naturais, energéticos ou matérias-primas, considerando processos biológicos, químicos ou físicos neles envolvidos.

H9 – Compreender a importância dos ciclos biogeoquímicos ou do fluxo energia para a vida, ou da ação de agentes ou fenômenos que podem causar alterações nesses processos.

H10 – Analisar perturbações ambientais, identificando fontes, transporte e(ou) destino dos poluentes ou prevendo efeitos em sistemas naturais, produtivos ou sociais.

H11 – Reconhecer benefícios, limitações e aspectos éticos da biotecnologia, considerando estruturas e processos biológicos envolvidos em produtos biotecnológicos.

H12 – Avaliar impactos em ambientes naturais decorrentes de atividades sociais ou econômicas, considerando interesses contraditórios.

Competência de área 4 – Compreender interações entre organismos e ambiente, em particular aquelas relacionadas à saúde humana, relacionando conhecimentos científicos, aspectos culturais e características individuais.

H13 – Reconhecer mecanismos de transmissão da vida, prevendo ou explicando a manifestação de características dos seres vivos.

H14 – Identificar padrões em fenômenos e processos vitais dos organismos, como manutenção do equilíbrio interno, defesa, relações com o ambiente, sexualidade, entre outros.

H15 – Interpretar modelos e experimentos para explicar fenômenos ou processos biológicos em qualquer nível de organização dos sistemas biológicos.

H16 – Compreender o papel da evolução na produção de padrões, processos biológicos ou na organização taxonômica dos seres vivos.

Competência de área 5 – Entender métodos e procedimentos próprios das ciências naturais e aplicá-los em diferentes contextos.

H17 – Relacionar informações apresentadas em diferentes formas de linguagem e representação usadas nas ciências físicas, químicas ou biológicas, como texto discursivo, gráficos, tabelas, relações matemáticas ou linguagem simbólica.

H18 – Relacionar propriedades físicas, químicas ou biológicas de produtos, sistemas ou procedimentos tecnológicos às finalidades a que se destinam.

H19 – Avaliar métodos, processos ou procedimentos das ciências naturais que contribuam para diagnosticar ou solucionar problemas de ordem social, econômica ou ambiental.

Competência de área 6 – Apropriar-se de conhecimentos da física para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas.

H20 – Caracterizar causas ou efeitos dos movimentos de partículas, substâncias, objetos ou corpos celestes.

H21 – Utilizar leis físicas e (ou) químicas para interpretar processos naturais ou tecnológicos inseridos no contexto da termodinâmica e(ou) do eletromagnetismo.

H22 – Compreender fenômenos decorrentes da interação entre a radiação e a matéria em suas manifestações em processos naturais ou tecnológicos, ou em suas implicações biológicas, sociais, econômicas ou ambientais.

H23 – Avaliar possibilidades de geração, uso ou transformação de energia em ambientes específicos, considerando implicações éticas, ambientais, sociais e/ou econômicas.

Competência de área 7 – Apropriar-se de conhecimentos da química para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas.

H24 – Utilizar códigos e nomenclatura da química para caracterizar materiais, substâncias ou transformações químicas.

H25 – Caracterizar materiais ou substâncias, identificando etapas, rendimentos ou implicações biológicas, sociais, econômicas ou ambientais de sua obtenção ou produção.

H26 – Avaliar implicações sociais, ambientais e/ou econômicas na produção ou no consumo de recursos energéticos ou minerais, identificando transformações químicas ou de energia envolvidas nesses processos.

H27 – Avaliar propostas de intervenção no meio ambiente aplicando conhecimentos químicos, observando riscos ou benefícios.

Competência de área 8 – Apropriar-se de conhecimentos da biologia para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas.

H28 – Associar características adaptativas dos organismos com seu modo de vida ou com seus limites de distribuição em diferentes ambientes, em especial em ambientes brasileiros.

H29 – Interpretar experimentos ou técnicas que utilizam seres vivos, analisando implicações para o ambiente, a saúde, a produção de alimentos, matérias primas ou produtos industriais.

H30 – Avaliar propostas de alcance individual ou coletivo, identificando aquelas que visam à preservação e a implementação da saúde individual, coletiva ou do ambiente.

Matriz de Referência de Ciências Humanas e suas Tecnologias

Competência de área 1 - Compreender os elementos culturais que constituem as identidades

H1 - Interpretar historicamente e/ou geograficamente fontes documentais acerca de aspectos da cultura.

H2 - Analisar a produção da memória pelas sociedades humanas.

H3 - Associar as manifestações culturais do presente aos seus processos históricos.

H4 - Comparar pontos de vista expressos em diferentes fontes sobre determinado aspecto da cultura.

H5 - Identificar as manifestações ou representações da diversidade do patrimônio cultural e artístico em diferentes sociedades.

Competência de área 2 - Compreender as transformações dos espaços geográficos como produto das relações socioeconômicas e culturais de poder.

H6 - Interpretar diferentes representações gráficas e cartográficas dos espaços geográficos.

H7 - Identificar os significados histórico-geográficos das relações de poder entre as nações

H8 - Analisar a ação dos estados nacionais no que se refere à dinâmica dos fluxos populacionais e no enfrentamento de problemas de ordem econômico-social.

H9 - Comparar o significado histórico-geográfico das organizações políticas e socioeconômicas em escala local, regional ou mundial.

H10 - Reconhecer a dinâmica da organização dos movimentos sociais e a importância da participação da coletividade na transformação da realidade histórico-geográfica.

Competência de área 3 - Compreender a produção e o papel histórico das instituições sociais, políticas e econômicas, associando-as aos diferentes grupos, conflitos e movimentos sociais.

H11 - Identificar registros de práticas de grupos sociais no tempo e no espaço.

H12 - Analisar o papel da justiça como instituição na organização das sociedades.

H13 - Analisar a atuação dos movimentos sociais que contribuíram para mudanças ou

rupturas em processos de disputa pelo poder.

H14 - Comparar diferentes pontos de vista, presentes em textos analíticos e interpretativos, sobre situação ou fatos de natureza histórico-geográfica acerca das instituições sociais, políticas e econômicas.

H15 - Avaliar criticamente conflitos culturais, sociais, políticos, econômicos ou ambientais ao longo da história.

Competência de área 4 - Entender as transformações técnicas e tecnológicas e seu impacto nos processos de produção, no desenvolvimento do conhecimento e na vida social.

H16 - Identificar registros sobre o papel das técnicas e tecnologias na organização do trabalho e/ou da vida social.

H17 - Analisar fatores que explicam o impacto das novas tecnologias no processo de territorialização da produção.

H18 - Analisar diferentes processos de produção ou circulação de riquezas e suas implicações sócio-espaciais.

H19 - Reconhecer as transformações técnicas e tecnológicas que determinam as várias formas de uso e apropriação dos espaços rural e urbano.

H20 - Selecionar argumentos favoráveis ou contrários às modificações impostas pelas novas tecnologias à vida social e ao mundo do trabalho.

Competência de área 5 - Utilizar os conhecimentos históricos para compreender e valorizar os fundamentos da cidadania e da democracia, favorecendo uma atuação consciente do indivíduo na sociedade.

H21 - Identificar o papel dos meios de comunicação na construção da vida social.

H22 - Analisar as lutas sociais e conquistas obtidas no que se refere às mudanças nas legislações ou nas políticas públicas.

H23 - Analisar a importância dos valores éticos na estruturação política das sociedades.

H24 - Relacionar cidadania e democracia na organização das sociedades.

H25 – Identificar estratégias que promovam formas de inclusão social.

Competência de área 6 - Compreender a sociedade e a natureza, reconhecendo suas interações no espaço em diferentes contextos históricos e geográficos.

H26 - Identificar em fontes diversas o processo de ocupação dos meios físicos e as relações da vida humana com a paisagem.

H27 - Analisar de maneira crítica as interações da sociedade com o meio físico, levando em consideração aspectos históricos e(ou) geográficos.

H28 - Relacionar o uso das tecnologias com os impactos sócio-ambientais em diferentes contextos histórico-geográficos.

H29 - Reconhecer a função dos recursos naturais na produção do espaço geográfico, relacionando-os com as mudanças provocadas pelas ações humanas.

H30 - Avaliar as relações entre preservação e degradação da vida no planeta nas diferentes escalas.

ANEXO

Objetos de conhecimento associados às Matrizes de Referência

- **Estudo dos aspectos linguísticos em diferentes textos: recursos expressivos da língua, procedimentos de construção e recepção de textos** - organização da macroestrutura semântica e a articulação entre idéias e proposições (relações lógico-semânticas).
- **Estudo do texto argumentativo, seus gêneros e recursos linguísticos: argumentação: tipo, gêneros e usos em língua portuguesa** - formas de apresentação de diferentes pontos de vista; organização e progressão textual; papéis sociais e comunicativos dos interlocutores, relação entre usos e propósitos comunicativos, função sociocomunicativa do gênero, aspectos da dimensão espaço-temporal em que se produz o texto.
- **Estudo dos aspectos linguísticos da língua portuguesa: usos da língua: norma culta e variação lingüística** - uso dos recursos linguísticos em relação ao contexto em que o texto é constituído: elementos de referência pessoal, temporal, espacial, registro linguístico, grau de formalidade, seleção lexical, tempos e modos verbais; uso dos recursos linguísticos em processo de coesão textual: elementos de articulação das sequências dos textos ou à construção da micro estrutura do texto.
- **Estudo dos gêneros digitais: tecnologia da comunicação e informação: impacto e função social** - o texto literário típico da cultura de massa: o suporte textual em gêneros digitais; a caracterização dos interlocutores na comunicação tecnológica; os recursos linguísticos e os gêneros digitais; a função social das novas tecnologias.

1. Linguagem, Códigos e suas Tecnologias

- **Estudo do texto: as sequências discursivas e os gêneros textuais no sistema de comunicação e informação** - modos de organização da composição textual; atividades de produção escrita e de leitura de textos gerados nas diferentes esferas sociais - públicas e privadas.

- **Estudo das práticas corporais: a linguagem corporal como integradora social e formadora de identidade** *performance* corporal e identidades juvenis; possibilidades de vivência crítica e emancipada do lazer; mitos e verdades sobre os corpos masculino e feminino na sociedade atual; exercício físico e saúde; o corpo e a expressão artística e cultural; o corpo no mundo dos símbolos e como produção da cultura; práticas corporais e autonomia; condicionamentos e esforços físicos; o esporte; a dança; as lutas; os jogos; as brincadeiras.

- **Produção e recepção de textos artísticos: interpretação e representação do mundo para o fortalecimento dos processos de identidade e cidadania** - Artes Visuais: estrutura morfológica, sintática, o contexto da obra artística, o contexto da comunidade. Teatro: estrutura morfológica, sintática, o contexto da obra artística, o contexto da comunidade, as fontes de criação. Música: estrutura morfológica, sintática, o contexto da obra artística, o contexto da comunidade, as fontes de criação. Dança: estrutura morfológica, sintática, o contexto da obra artística, o contexto da comunidade, as fontes de criação. Conteúdos estruturantes das linguagens artísticas (Artes Visuais, Dança, Música, Teatro), elaborados a partir de suas estruturas morfológicas e sintáticas; inclusão, diversidade e multiculturalidade: a valorização da pluralidade expressada nas produções estéticas e artísticas das minorias sociais e dos portadores de necessidades especiais educacionais.

- **Estudo do texto literário: relações entre produção literária e processo social, concepções artísticas, procedimentos de construção e recepção de textos** - produção literária e processo social; processos de formação literária e de formação nacional; produção de textos literários, sua recepção e a constituição do patrimônio literário nacional; relações entre a dialética cosmopolitismo/localismo e a produção literária nacional; elementos de continuidade e ruptura entre os diversos momentos da literatura brasileira; associações entre concepções artísticas e procedimentos de construção do texto literário em seus gêneros (épico/narrativo, lírico e dramático) e formas diversas.; articulações entre os recursos expressivos e estruturais do texto literário e o processo social relacionado ao momento de sua produção; representação literária: natureza, função, organização e estrutura do texto literário; relações entre literatura, outras artes e outros saberes.

2. 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.

3. Ciências da Natureza e suas Tecnologias

3.1 Física

- **Conhecimentos básicos e fundamentais** - Noções de ordem de grandeza. Notação Científica. Sistema Internacional de Unidades. Metodologia de investigação: a procura de regularidades e de sinais na interpretação física do mundo. Observações e mensurações: representação de grandezas físicas como grandezas mensuráveis. Ferramentas básicas: gráficos e vetores. Conceituação de grandezas vetoriais e escalares. Operações básicas com vetores.
- **O movimento, o equilíbrio e a descoberta de leis físicas** - Grandezas fundamentais da mecânica: tempo, espaço, velocidade e aceleração. Relação histórica entre força e movimento. Descrições do movimento e sua interpretação: quantificação do movimento e sua descrição matemática e gráfica. Casos especiais de movimentos e suas regularidades observáveis. Conceito de inércia. Noção de sistemas de referência inerciais e não inerciais. Noção dinâmica de massa e quantidade de movimento (momento linear). Força e variação da quantidade de movimento. Leis de Newton. Centro de massa e a idéia de pontomaterial. Conceito de forças externas e internas. Lei da conservação da quantidade de movimento (momento linear) e teorema do impulso. Momento de uma força (torque). Condições de equilíbrio estático de ponto material e de corpos rígidos. Força de atrito, força peso, força normal de contato e tração. Diagramas de forças. Identificação das forças que atuam nos movimentos circulares. Noção de força centrípeta e sua quantificação. A hidrostática: aspectos históricos e variáveis relevantes. Empuxo. Princípios de Pascal, Arquimedes e Stevin condições de flutuação, relação entre diferença de nível e pressão hidrostática.
- **Energia, trabalho e potência** - Conceituação de trabalho, energia e potência. Conceito de energia potencial e de energia cinética. Conservação de energia mecânica e dissipação de energia. Trabalho força gravitacional e energia potencial gravitacional. Forças conservativas e dissipativas.
- **A Mecânica e o funcionamento do Universo** - Força peso. Aceleração gravitacional. Lei da Gravitação Universal. Leis de Kepler. Movimentos de corpos celestes. Influência na Terra: marés e variações climáticas. Concepções históricas sobre a origem do universo e sua evolução.
- **Fenômenos Elétricos e Magnéticos** - Carga elétrica e corrente elétrica. Lei de Coulomb. Campo elétrico e potencial elétrico. Linhas de campo. Superfícies equipotenciais. Poder das pontas. Blindagem. Capacitores. Efeito Joule. Lei de Ohm. Resistência elétrica e resistividade. Relações entre grandezas elétricas: tensão, corrente, potência e energia. Circuitos elétricos simples. Correntes contínua

e alternada. Medidores elétricos. Representação gráfica de circuitos. Símbolos convencionais. Potência e consumo de energia em dispositivos elétricos. Campo magnético. Ímãs permanentes. Linhas de campo magnético. Campo magnético terrestre.

- **Oscilações, ondas, óptica e radiação** - Feixes e frentes de ondas. Reflexão e refração. Óptica geométrica: lentes e espelhos. Formação de imagens. Instrumentos ópticos simples. Fenômenos ondulatórios. Pulsos e ondas. Período, frequência, ciclo. Propagação: relação entre velocidade, frequência e comprimento de onda. Ondas em diferentes meios de propagação.
- **O calor e os fenômenos térmicos** - Conceitos de calor e de temperatura. Escalas termométricas. Transferência de calor e equilíbrio térmico. Capacidade calorífica e calor específico. Condução do calor. Dilatação térmica. Mudanças de estado físico e calor latente de transformação. Comportamento de Gases ideais. Máquinas térmicas. Ciclo de Carnot. Leis da Termodinâmica. Aplicações e fenômenos térmicos de uso cotidiano. Compreensão de fenômenos climáticos relacionados ao ciclo da água.

3.2 Química

- **Transformações Químicas** - Evidências de transformações químicas. Interpretando transformações químicas. Sistemas Gasosos: Lei dos gases. Equação geral dos gases ideais, Princípio de Avogadro, conceito de molécula; massa molar, volume molar dos gases. Teoria cinética dos gases. Misturas gasosas. Modelo corpuscular da matéria. Modelo atômico de Dalton. Natureza elétrica da matéria: Modelo Atômico de Thomson, Rutherford, Rutherford-Bohr. Átomos e sua estrutura. Número atômico, número de massa, isótopos, massa atômica. Elementos químicos e Tabela Periódica. Reações químicas.
- **Representação das transformações químicas** - Fórmulas químicas. Balanceamento de equações químicas. Aspectos quantitativos das transformações químicas. Leis ponderais das reações químicas. Determinação de fórmulas químicas. Grandezas Químicas: massa, volume, mol, massa molar, constante de Avogadro. Cálculos estequiométricos.
- **Materiais, suas propriedades e usos** - Propriedades de materiais. Estados físicos de materiais. Mudanças de estado. Misturas: tipos e métodos de separação. Substâncias químicas: classificação e características gerais. Metais e Ligas metálicas. Ferro, cobre e alumínio. Ligações metálicas. Substâncias iônicas: características e propriedades. Substâncias iônicas do grupo: cloreto, carbonato, nitrato e sulfato. Ligação iônica. Substâncias moleculares: características e propriedades. Substâncias moleculares: H_2 , O_2 , N_2 , Cl_2 , NH_3 , H_2O , HCl , CH_4 . Ligação Covalente. Polaridade de moléculas. Forças intermoleculares. Relação entre estruturas, propriedade e aplicação das substâncias.

- **Água** - Ocorrência e importância na vida animal e vegetal. Ligação, estrutura e propriedades. Sistemas em Solução Aquosa: Soluções verdadeiras, soluções coloidais e suspensões. Solubilidade. Concentração das soluções. Aspectos qualitativos das propriedades coligativas das soluções. Ácidos, Bases, Sais e Óxidos: definição, classificação, propriedades, formulação e nomenclatura. Conceitos de ácidos e base. Principais propriedades dos ácidos e bases: indicadores, condutibilidade elétrica, reação com metais, reação de neutralização.
- **Transformações Químicas e Energia** - Transformações químicas e energia calorífica. Calor de reação. Entalpia. Equações termoquímicas. Lei de Hess. Transformações químicas e energia elétrica. Reação de oxirredução. Potenciais padrão de redução. Pilha. Eletrólise. Leis de Faraday. Transformações nucleares. Conceitos fundamentais da radioatividade. Reações de fissão e fusão nuclear. Desintegração radioativa e radioisótopos.
- **Dinâmica das Transformações Químicas** - Transformações Químicas e velocidade. Velocidade de reação. Energia de ativação. Fatores que alteram a velocidade de reação: concentração, pressão, temperatura e catalisador.
- **Transformação Química e Equilíbrio** - Caracterização do sistema em equilíbrio. Constante de equilíbrio. Produto iônico da água, equilíbrio ácido-base e pH. Solubilidade dos sais e hidrólise. Fatores que alteram o sistema em equilíbrio. Aplicação da velocidade e do equilíbrio químico no cotidiano.
- **Compostos de Carbono** - Características gerais dos compostos orgânicos. Principais funções orgânicas. Estrutura e propriedades de Hidrocarbonetos. Estrutura e propriedades de compostos orgânicos oxigenados. Fermentação. Estrutura e propriedades de compostos orgânicos nitrogenados. Macromoléculas naturais e sintéticas. Noções básicas sobre polímeros. Amido, glicogênio e celulose. Borracha natural e sintética. Polietileno, poliestireno, PVC, Teflon, náilon. Óleos e gorduras, sabões e detergentes sintéticos. Proteínas e enzimas.
- **Relações da Química com as Tecnologias, a Sociedade e o Meio Ambiente** - Química no cotidiano. Química na agricultura e na saúde. Química nos alimentos. Química e ambiente. Aspectos científico-tecnológicos, socioeconômicos e ambientais associados à obtenção ou produção de substâncias químicas. Indústria Química: obtenção e utilização do cloro, hidróxido de sódio, ácido sulfúrico, amônia e ácido nítrico. Mineração e Metalurgia. Poluição e tratamento de água. Poluição atmosférica. Contaminação e proteção do ambiente.
- **Energias Químicas no Cotidiano** - Petróleo, gás natural e carvão. Madeira e hulha. Biomassa. Biocombustíveis. Impactos ambientais de combustíveis fósseis. Energia nuclear. Lixo atômico. Vantagens e desvantagens do uso de energia nuclear.

de desenvolvimento humano. Principais doenças que afetam a população brasileira: caracterização, prevenção e profilaxia. Noções de primeiros socorros. Doenças sexualmente transmissíveis. Aspectos sociais da biologia: uso indevido de drogas; gravidez na adolescência; obesidade. Violência e segurança pública. Exercícios físicos e vida saudável. Aspectos biológicos do desenvolvimento sustentável. Legislação e cidadania.

4. Ciências Humanas e suas Tecnologias

□ **Diversidade cultural, conflitos e vida em sociedade**

- Cultura Material e imaterial; patrimônio e diversidade cultural no Brasil.
- A Conquista da América. Conflitos entre europeus e indígenas na América colonial. A escravidão e formas de resistência indígena e africana na América.
- História cultural dos povos africanos. A luta dos negros no Brasil e o negro na formação da sociedade brasileira.
- História dos povos indígenas e a formação sócio-cultural brasileira.
- Movimentos culturais no mundo ocidental e seus impactos na vida política e social.

□ **Formas de organização social, movimentos sociais, pensamento político e ação do Estado**

- Cidadania e democracia na Antiguidade; Estado e direitos do cidadão a partir da Idade Moderna; democracia direta, indireta e representativa.
- Revoluções sociais e políticas na Europa Moderna.
- Formação territorial brasileira; as regiões brasileiras; políticas de reordenamento territorial.
- As lutas pela conquista da independência política das colônias da América.
- Grupos sociais em conflito no Brasil imperial e a construção da nação.
- O desenvolvimento do pensamento liberal na sociedade capitalista e seus críticos nos séculos XIX e XX.
- Políticas de colonização, migração, imigração e emigração no Brasil nos séculos XIX e XX.
- A atuação dos grupos sociais e os grandes processos revolucionários do século XX: Revolução Bolchevique, Revolução Chinesa, Revolução Cubana.
- Geopolítica e conflitos entre os séculos XIX e XX: Imperialismo, a ocupação da Ásia e da África, as Guerras Mundiais e a Guerra Fria.
- Os sistemas totalitários na Europa do século XX: nazi-fascista, franquismo, salazarismo e stalinismo. Ditaduras políticas na América Latina: Estado Novo no Brasil e ditaduras na América.
- Conflitos político-culturais pós-Guerra Fria, reorganização política internacional e

3.3 Biologia

- **Moléculas, células e tecidos** - Estrutura e fisiologia celular: membrana, citoplasma e núcleo. Divisão celular. Aspectos bioquímicos das estruturas celulares. Aspectos gerais do metabolismo celular. Metabolismo energético: fotossíntese e respiração. Codificação da informação genética. Síntese protéica. Diferenciação celular. Principais tecidos animais e vegetais. Origem e evolução das células. Noções sobre células-tronco, clonagem e tecnologia do DNA recombinante. Aplicações de biotecnologia na produção de alimentos, fármacos e componentes biológicos. Aplicações de tecnologias relacionadas ao DNA a investigações científicas, determinação da paternidade, investigação criminal e identificação de indivíduos. Aspectos éticos relacionados ao desenvolvimento biotecnológico. Biotecnologia e sustentabilidade.
- **Hereditariedade e diversidade da vida** - Princípios básicos que regem a transmissão de características hereditárias. Concepções pré-mendelianas sobre a hereditariedade. Aspectos genéticos do funcionamento do corpo humano. Antígenos e anticorpos. Grupos sanguíneos, transplantes e doenças auto-imunes. Neoplasias e a influência de fatores ambientais. Mutações gênicas e cromossômicas. Aconselhamento genético. Fundamentos genéticos da evolução. Aspectos genéticos da formação e manutenção da diversidade biológica.
- **Identidade dos seres vivos** - Níveis de organização dos seres vivos. Vírus, procariontes e eucariontes. Autótrofos e heterótrofos. Seres unicelulares e pluricelulares. Sistemática e as grandes linhas da evolução dos seres vivos. Tipos de ciclo de vida. Evolução e padrões anatômicos e fisiológicos observados nos seres vivos. Funções vitais dos seres vivos e sua relação com a adaptação desses organismos a diferentes ambientes. Embriologia, anatomia e fisiologia humana. Evolução humana. Biotecnologia e sistemática.
- **Ecologia e ciências ambientais** - Ecossistemas. Fatores bióticos e abióticos. Habitat e nicho ecológico. A comunidade biológica: teia alimentar, sucessão e comunidade clímax. Dinâmica de populações. Interações entre os seres vivos. Ciclos biogeoquímicos. Fluxo de energia no ecossistema. Biogeografia. Biomas brasileiros. Exploração e uso de recursos naturais. Problemas ambientais: mudanças climáticas, efeito estufa; desmatamento; erosão; poluição da água, do solo e do ar. Conservação e recuperação de ecossistemas. Conservação da biodiversidade. Tecnologias ambientais. Noções de saneamento básico. Noções de legislação ambiental: água, florestas, unidades de conservação; biodiversidade.
- **Origem e evolução da vida** - A biologia como ciência: história, métodos, técnicas e experimentação. Hipóteses sobre a origem do Universo, da Terra e dos seres vivos. Teorias de evolução. Explicações pré-darwinistas para a modificação das espécies. A teoria evolutiva de Charles Darwin. Teoria sintética da evolução. Seleção artificial e seu impacto sobre ambientes naturais e sobre populações humanas.
- **Qualidade de vida das populações humanas** - Aspectos biológicos da pobreza e do desenvolvimento humano. Indicadores sociais, ambientais e econômicos. Índice

os organismos multilaterais nos séculos XX e XXI.

- A luta pela conquista de direitos pelos cidadãos: direitos civis, humanos, políticos e sociais. Direitos sociais nas constituições brasileiras. Políticas afirmativas.
- Vida urbana: redes e hierarquia nas cidades, pobreza e segregação espacial.

□ **Características e transformações das estruturas produtivas**

- Diferentes formas de organização da produção: escravismo antigo, feudalismo, capitalismo, socialismo e suas diferentes experiências.
- Economia agro-exportadora brasileira: complexo açucareiro; a mineração no período colonial; a economia cafeeira; a borracha na Amazônia.
- Revolução Industrial: criação do sistema de fábrica na Europa e transformações no processo de produção. Formação do espaço urbano-industrial. Transformações na estrutura produtiva no século XX: o fordismo, o toyotismo, as novas técnicas de produção e seus impactos.
- A industrialização brasileira, a urbanização e as transformações sociais e trabalhistas.
- A globalização e as novas tecnologias de telecomunicação e suas conseqüências econômicas, políticas e sociais.
- Produção e transformação dos espaços agrários. Modernização da agricultura e estruturas agrárias tradicionais. O agronegócio, a agricultura familiar, os assalariados do campo e as lutas sociais no campo. A relação campo-cidade.

□ **Os domínios naturais e a relação do ser humano com o ambiente**

- Relação homem-natureza, a apropriação dos recursos naturais pelas sociedades ao longo do tempo. Impacto ambiental das atividades econômicas no Brasil. Recursos minerais e energéticos: exploração e impactos. Recursos hídricos; bacias hidrográficas e seus aproveitamentos.
- As questões ambientais contemporâneas: mudança climática, ilhas de calor, efeito estufa, chuva ácida, a destruição da camada de ozônio. A nova ordem ambiental internacional; políticas territoriais ambientais; uso e conservação dos recursos naturais, unidades de conservação, corredores ecológicos, zoneamento ecológico e econômico.
- Origem e evolução do conceito de sustentabilidade.
- Estrutura interna da terra. Estruturas do solo e do relevo; agentes internos e externos modeladores do relevo.
- Situação geral da atmosfera e classificação climática. As características

climáticas do território brasileiro.

- Os grandes domínios da vegetação no Brasil e no mundo.

□ **Representação espacial**

- Projeções cartográficas; leitura de mapas temáticos, físicos e políticos; tecnologias modernas aplicadas à cartografia.

Apêndice 1

Questionário de Física Térmica – Versão em Português

Este questionário investiga os seus conhecimentos sobre os conceitos de calor e temperatura.

- Para ajudá-lo a visualizar cada situação, pense num grupo de amigos que estão na cozinha de uma cantina. Imagine que eles estão atentos e interessados em compreender os fenômenos corriqueiros que ocorrem neste ambiente.
- Para cada questão escolha a resposta que está mais próxima de seu entendimento.
- Seja cuidadoso ao marcar a alternativa que deseja. Algumas questões possuem cinco alternativas.

Questões

1. Qual é a temperatura mais provável do cubos de gelo que se encontram armazenados no congelador de um freezer?
a. -10°C
b. 0°C
c. 5°C
d. A temperatura depende do tamanho dos cubos de gelo.
2. Francisco pega seis cubos de gelo no freezer e coloca quatro deles dentro de um copo com água, deixando os outros dois sobre o tampo de um balcão. Ele agita várias vezes o copo até que as pedras de gelo estejam muito pequenas e tenham parado de derreter. Qual é a temperatura mais provável da água nesta situação?
a. -10°C b. 0°C c. 5°C d. 10°C
3. Os cubos de gelo que Francisco deixou sobre o balcão quase derreteram por completo e repousam sobre uma poça de água. Qual é a temperatura mais provável desses cubos de gelo menores?
a. -10°C b. 0°C c. 5°C d. 10°C
4. Sobre o fogão está uma chaleira cheia de água. A água começou a ferver rapidamente. A temperatura mais provável da água é:
a. 88°C
b. 98°C
c. 110°C
d. Nenhuma das respostas acima podem estar corretas.
5. Cinco minutos mais tarde, a água na chaleira ainda ferve. A temperatura mais provável da água agora é:
a. 88°C b. 98°C c. 110°C d. 120°C
6. Qual você acha que é a temperatura do vapor formado, acima da água fervente, na chaleira?
a. 88°C b. 98°C
c. 110°C d. 120°C
7. Simone pega dois copos de água à 40°C e os mistura a um copo de água a 10°C . Qual é a temperatura mais provável da mistura?
a. 20°C b. 25°C
c. 30° d. 50°C
8. ue deve usar água fervendo para fazer uma xícara de chá. Ele diz para seus amigos: “Eu não poderia fazer chá se estivesse acampando numa montanha alta, porque a água não ferve em grandes altitudes.”
 - a. Junior diz: “Sim, ela ferve, mas a água fervente não é tão quente como aqui”.
 - b. Talita diz: “Não é verdade. A água sempre ferve na mesma temperatura”.
 - c. Luiz diz: “O ponto de fusão da água diminui, mas a água ainda está a 100 graus”.
 - d. Samantha diz: “Eu concordo com Raphael. A água nunca atingirá seu ponto de ebulição”.Com quem você concorda?

9. Pedro pega uma lata de coca-cola e uma garrafa plástica de coca-cola no refrigerador, onde elas haviam sido colocadas na noite anterior. Ele rapidamente coloca um termômetro na coca-cola contida na lata. A temperatura é de 7°C. Qual é a temperatura mais provável da garrafa de plástico e da coca-cola em seu interior?
- Elas estão ambas abaixo de 7°C.
 - Elas estão ambas a 7°C.
 - Elas estão ambas acima de 7°C.
 - A coca-cola está a 7°C, mas a garrafa está a mais de 7°C.
 - Depende da quantidade de coca-cola e/ou do tamanho da garrafa.
10. Poucos minutos mais tarde, Cassiano levanta a lata de coca-cola e diz para todos que superfície do balcão que estava embaixo da lata está mais fria que o resto do balcão.
- Gabriel diz: "O frio foi transferido da coca-cola para o balcão."
 - Alan diz: "Não há energia deixada no balcão pela lata."
 - Sabrina diz: "Um pouco de calor foi transferido do balcão para coca-cola."
 - Edilaine diz: "A lata faz com que o calor embaixo dela se desloque através dela para o balcão." Qual das explicações você acha que é a melhor?
11. Pâmela pergunta para o grupo de amigos: "Se eu colocar 100 gramas de gelo a 0°C e 100 gramas de água a 0°C dentro do freezer, qual dos dois perderá a maior quantidade de calor?"
- Cristiane diz: "Os 100 gramas de gelo."
 - Bernardo diz: "Os 100 gramas de água."
 - Marcelo diz: "Nenhum dos dois, porque ambos contêm a mesma quantidade de calor."
 - Mateus diz: "Não há resposta, porque gelo não contém nenhum calor."
 - Julia diz: "Não há resposta, porque você não pode ter água a 0°C." Com qual dos amigos de Pâmela você concorda?
12. Melissa está fervendo água numa panela sobre o fogão. O que você acha que está dentro das bolhas que se formam na água fervente? Principalmente:
- Ar
 - Gás oxigênio e hidrogênio
 - Vapor de água
 - Não há nada nas bolhas
13. Depois de cozinhar alguns ovos na água fervente, Melissa esfria os ovos colocando-os numa vasilha de água fria. Qual das seguintes afirmações explica o processo de esfriamento?
- A temperatura é transferida do ovo para a água.
 - O frio se move da água para os ovos.
 - Objetos quentes esfriam naturalmente.
 - A energia é transferida dos ovos para a água.
14. Janete diz que não gosta de sentar em cadeiras de metal na sala porque "elas são mais frias do que as de plástico."
- Rafael concorda e diz: "Elas são mais frias porque o metal é naturalmente mais frio do que o plástico."
 - Mário diz: "Elas não são mais frias, elas estão com a mesma temperatura."
 - Luiz diz: "Elas não são mais frias, as de metal só parecem mais frias porque são mais pesadas."
 - Samantha diz: "Elas são mais frias porque o metal tem menos calor para perder do que o plástico." Quem você acha que está correto?
15. Um grupo está ouvindo a previsão do tempo no rádio. Eles ouvem: "... à noite fará uns 5°C, mais frio do que os 10°C da noite passada."
- Marta diz: "Isso significa que esta noite será duas vezes mais fria que a noite passada."
 - Alice diz: "Não está certo. 5°C não é duas vezes mais frio que 10°C."
 - Rui diz: "Está parcialmente correto, mas ela deveria dizer que 10°C é duas vezes mais quente do que 5°C."
 - Gabi diz: "Está parcialmente correto, mas ela deveria ter dito que 5°C é a metade do frio de 10°C. Com qual afirmação você mais concorda?"

16. Tom pega uma régua de metal e uma de madeira de seu estojo. Ele diz que sente a régua de metal mais fria que a de madeira. Qual é sua explicação preferida?
- O metal retira energia da sua mão mais rapidamente que a madeira.
 - A madeira é naturalmente uma substância mais quente do que o metal.
 - A régua de madeira contém mais calor que a régua de metal.
 - Metais são melhores irradiadores de calor do que a madeira.
 - O frio flui mais facilmente de um meta.
17. Alzira pegou duas garrafas de vidro contendo água a 20°C e as embrulhou em toalhas de banho. Uma das toalhas estava molhada e a outra estava seca. Vinte minutos mais tarde, ela mede a temperatura da água em cada garrafa. A água na garrafa embrulhada na toalha molhada estava a 18°C e a embrulhada na toalha seca estava a 22°C. A temperatura mais provável da sala durante esse experimento era:
- 26°C
 - 21°C
 - 20°C
 - 18°C
18. Daniel pega simultaneamente [ao mesmo tempo?] duas embalagens de leite achocolatado, uma fria, retirada da geladeira e outra morna, que havia sido deixada sobre o tampo do balcão por algum tempo. Porque você acha que a embalagem retirada da geladeira parece mais fria que a do balcão? Comparada à embalagem morna, a embalagem fria:
- contém mais frio.
 - contém menos calor.
 - é pior condutora de calor.
 - retira calor mais rapidamente da mão de Dan.
 - cede calor mais rapidamente para a mão de Dan.
19. Ronaldo sabe que sua mãe faz sopa na panela de pressão porque ela cozinha mais rápido do que uma panela normal, mas ele não sabe o porquê disso. [Painelas de pressão têm a tampa selada, o que faz com que a pressão dentro delas atinja um valor maior que o da pressão atmosférica.]
- Amanda diz: “Isso acontece porque a pressão faz a água ferver acima de 100°C.”
 - Carlos diz: “Isso ocorre porque a pressão alta gera calor extra.”
 - Fernanda diz: “É porque o vapor está numa temperatura maior do que a sopa fervente.”
 - Tomaz diz: “É porque panelas de pressão propagam o calor mais uniformemente através dos alimentos.”
- Com qual deles você mais concorda?
20. Patrícia acredita que seu pai faz bolo na prateleira superior do forno elétrico porque a parte de cima é mais quente que a de baixo.
- Pâmela diz que é mais quente porque o calor sobe.
 - Samuel diz que é mais quente porque bandejas de metal concentram mais calor.
 - Rubens diz é mais quente na parte superior do forno porque quanto mais quente o ar menos denso ele é.
 - Tiago não concorda com eles e diz que não é possível que a parte superior do forno seja mais quente. Qual deles você acha que está correto?
21. Beatriz está lendo uma questão de múltipla escolha de um livro texto: “Suar te esfria porque o suor sobre sua pele:
- umedece a superfície, e superfícies úmidas retiram mais calor do que superfícies secas.”
 - drena o calor dos poros e o espalha pela superfície da pele.”
 - está à mesma temperatura que tua pele, mas está evaporando e assim afastando o calor dela.”
 - é ligeiramente mais frio do que tua pele devido à evaporação, e dessa forma o calor é transferido da tua pele para o suor.”
- Qual resposta você diria para ela escolher?
22. Quando André usa uma bomba para encher o pneu de sua bicicleta, ele observa que a bomba fica bem quente. Qual das explicações abaixo parece ser a melhor?

- a. Energia foi transferida para a bomba.
b. Temperatura foi transferida para a bomba.
c. Calor propagou-se das suas mãos para a bomba.
d. O metal da bomba provoca o aumento de temperatura.
23. Porque vestimos agasalhos quando está frio?
- a. Para manter o frio fora da roupa.
b. Para gerar calor.
c. Para reduzir a perda de calor.
d. Todas as três respostas acima estão corretas.
24. Vitor pega alguns picolés do congelador, onde ele os tinha colocado no dia anterior, e disse a todos que os palitos de madeira estão a uma temperatura mais alta do que a parte gelada (o picolé).
- a. Débora diz: “Você está certo porque os palitos de madeira não ficam tão frios quanto o picolé”.
b. Ian diz: “Você está certo porque o picolé contém mais frio que a parte de madeira.”
c. Márcio diz: “Você está errado, eles apenas parecem diferentes porque o palito contém mais calor.”
d. Ana diz: “Eu acho que eles estão à mesma temperatura porque estão juntos.”
Com qual deles você concorda?
25. Gabi está falando sobre um programa de televisão que ela viu na noite anterior. “Eu vi físicos fazerem supercondutores magnéticos a uma temperatura de -260°C .”
- a. Jorge duvida disso: “Você deve estar errada. Você não pode ter temperaturas tão baixas assim..”
b. Kátia discorda dele e diz: “Sim, você pode. Não existe um limite inferior para baixas temperaturas.”
c. Leonardo acredita que ele está certo: “Eu acho que o ímã estava perto da temperatura mais baixa possível.”
d. Gabi não está segura: “Eu acho que supercondutores são bons condutores de calor, logo você não pode resfriá-los a tão baixas temperaturas.
Quem você acha que está correto?”
26. Quatro estudantes estão discutindo coisas que fizeram quando crianças. A seguinte conversa foi ouvida: Patrícia: “Eu costumava envolver minhas bonecas em cobertores mas nunca entendi porque elas não esquentavam.”
- a. Marcelo afirma: “É porque os cobertores que você utilizava eram provavelmente isolantes térmicos ruins.
b. Laura replica: “É porque os cobertores que você usou eram provavelmente pobres condutores.”
c. Fabiana replica: “É porque as bonecas são feitas de um material que não mantém bem o calor.”
d. Cecília replica: “É porque as bonecas eram feitas de um material que levava muito tempo para se aquecer.”
e. Joana afirma: “Vocês estão todos errados.”
Com quem você concorda?

Apêndice 2

Relação Conceito/Item (Questionário de Física Térmica)

Conceito	Questão/item
1	10d, 21a, 21b, 22c,
2	21b, 22c,
3	10a, 13b, 15c (?), 18a, 18b, 23a, 24b, 24c
4	18c, 18e (?)
5	7b (?), 11c, 15a, 15c, 15d,
6	7a (?), 7d,
7	
8	
9	10b, 18a, 18b, 22c,
10	5d
11	8d, 19b (?), 19d,
12	7d (?), 10b, 11d, 26c,
13	1d, 9e, 14c,
14	25b
15	3c, 3d, 4a, 4c, 5c
16	20a,
17	20a,
18	10d, 13b,
19	13a, 22b,
20	1c, 2a, 6a, 9a, 9c, 9d, 17b, 17c, 17d, 24a, 24b
21	3a, 13c,
22	26b
23	
24	9d, 14a, 16b, 16c, 24a, 24b, 24c
25	9d, 14d, 16d, 16e, 20b, 22d,
26	25d
27	
28	4d, 8b, 8c, 19d,
29	1b,
30	2c, 2d, 11e,
31	6c, 6d, 19c,
32	23b, 26a,
33	26d
34	
35	12a, 12b, 12d

APÊNDICE 3

Objetos de conhecimento associados à matriz de referência de ciências da natureza e suas tecnologias

1. Conhecimentos Básicos e Fundamentais

- 1.1) Noções de ordem de grandeza;
- 1.2) Notação científica;
- 1.3) Sistema internacional de unidades;
- 1.4) Metodologia de investigação: a procura de regularidades e de sinais na interpretação física do mundo;
- 1.5) Observações e mensurações: representação de grandezas físicas como grandezas mensuráveis;
- 1.6) Ferramentas básicas: gráficos e vetores;
- 1.7) Conceituação de grandezas vetoriais e escalares;
- 1.8) Operações básicas com vetores.

2. O movimento, o equilíbrio e a descoberta de leis físicas

- 2.1) Grandezas fundamentais da mecânica: tempo, espaço, velocidade e aceleração;
- 2.2) Relação histórica entre força e movimento;
- 2.3) Descrições do movimento e sua interpretação: quantificação do movimento e sua descrição matemática e gráfica;
- 2.4) Casos especiais de movimentos e suas regularidades observáveis;
- 2.5) Conceito de inércia. Noção de sistemas de referência inerciais e não inerciais;
- 2.6) Noção dinâmica de massa e quantidade de movimento (momento linear);
- 2.7) Força e variação da quantidade de movimento;
- 2.8) Leis de Newton;
- 2.9) Centro de massa e a idéia de ponto material;
- 2.10) Conceito de forças externas e internas. Lei da conservação da quantidade de movimento (momento linear) e teorema do impulso;
- 2.11) Momento de uma força (torque);
- 2.12) Condições de equilíbrio estático de ponto material e de corpos rígidos;
- 2.13) Força de atrito, força peso, força normal de contato e tração;
- 2.14) Diagramas de forças;
- 2.15) Identificação das forças que atuam nos movimentos circulares;
- 2.16) Noção de força centrípeta e sua quantificação;
- 2.17) A hidrostática: aspectos históricos e variáveis relevantes;
- 2.18) Empuxo Princípios de Pascal, Arquimedes e Stevin: condições de flutuação, relação entre diferença de nível e pressão hidrostática;

3. Energia, trabalho e potência

- 3.1) Conceituação de trabalho, energia e potência;
- 3.2) Conceito de energia potencial e de energia cinética.
- 3.3) Conservação de energia mecânica e dissipação de energia;
- 3.4) Trabalho da força gravitacional e energia potencial gravitacional;
- 3.5) Forças conservativas e dissipativas;

4. A Mecânica e o funcionamento do Universo
 - 4.1) Força peso. Aceleração gravitacional;
 - 4.2) Lei da Gravitação Universal;
 - 4.3) Leis de Kepler;
 - 4.4) Movimentos de corpos celestes;
 - 4.5) Influência na Terra: marés e variações climáticas;
 - 4.6) Concepções históricas sobre a origem do universo e sua evolução;

5. Fenômenos Elétricos e Magnéticos
 - 5.1) Carga elétrica e corrente elétrica;
 - 5.2) Lei de Coulomb;
 - 5.3) Campo elétrico Linhas de campo;
 - 5.4) Potencial elétrico e Superfícies equipotenciais;
 - 5.5) Poder das pontas. Blindagem. Capacitores;
 - 5.6) Efeito Joule;
 - 5.7) Lei de Ohm;
 - 5.8) Resistência elétrica e resistividade;
 - 5.9) Relações entre grandezas elétricas: tensão, corrente, potência e energia;
 - 5.10) Circuitos elétricos simples;
 - 5.11) Correntes contínua e alternada;
 - 5.12) Medidores elétricos;
 - 5.13) Representação gráfica de circuitos. Símbolos convencionais;
 - 5.14) Potência e consumo de energia em dispositivos elétricos;
 - 5.15) Campo magnético. Ímãs permanentes. Linhas de campo magnético. Campo magnético terrestre;

6. Oscilações, ondas, óptica e radiação
 - 6.1) Feixes e frentes de ondas;
 - 6.2) Reflexão e refração;
 - 6.3) Óptica geométrica: lentes e espelhos. Formação de imagens
 - 6.4) Instrumentos ópticos simples;
 - 6.5) Fenômenos ondulatórios;
 - 6.6) Pulsos e ondas. Período, frequência, ciclo;
 - 6.7) Propagação: relação entre velocidade, frequência e comprimento de onda.
 - 6.8) Ondas em diferentes meios de propagação.

7. O calor e os fenômenos térmicos
 - 7.1) Conceitos de calor e de temperatura;
 - 7.2) Escalas termométricas;
 - 7.3) Transferência de calor e equilíbrio térmico;
 - 7.4) Capacidade calorífica e calor específico;
 - 7.5) Condução do calor;
 - 7.6) Dilatação térmica;
 - 7.7) Mudanças de estado físico e calor latente de transformação;
 - 7.8) Comportamento de Gases ideais;
 - 7.9) Máquinas térmicas. Ciclo de Carnot. Leis da Termodinâmica;
 - 7.10) Aplicações e fenômenos térmicos de uso cotidiano;
 - 7.11) Compreensão de fenômenos climáticos relacionados ao ciclo da água.

Apêndice 4

Levantamento de Dados - Enem 2009 a 2011

ENEM 2009/1

Item	Gabarito	Competência	Habilidade	Habilidades Possíveis	Tema de Física Abordado	Tema de Física Abordado	Extensão (linhas em 1/2 página)	Enunciado (linhas em 1/2 página)
1	c	3	9	9 ou 21	7	7.7	9	1
9	b	3	12	9, 12 ou 21	7	7.7	9	3
17	d	2	7		2	2.18	15	1
19	d	6	23		3	3.1	14	1
23	e	2	6		1	1.5	21	6
25	d	1	1		6	6.7	10	3
26	c	2	5		5	5.9	9	13
28	d	6	22	22 ou 23	7	7.3	8	1
31	c	5	17	17 ou 21	1	1.6	7	2
32	e	6	20		2	2.3	12	2
33	e	5	19	17, 19 ou 23	3	3.1	5	4
34	c	6	21		7	7.7	11	2
35	c	6	22		7	7.4	7	5
42	e	6	23		3	3.1	14	4
44	d	1	1		6	6.5	18	7
45	b	2	6		5	5.14	3	4

Item	Linguagem não textual	interpretação de imagens / texto	Nível de Contextualização	Há cobrança explícita de Conteúdo	Necessita de memorização de fórmula	Laboratório é diferencial	quanti / quali/ semi	conexão cognitiva (conteúdos)	conexão cognitiva (áreas)
1	não	não	2	sim	não	não	ql	não	não
9	não	não	0	sim	não	não	ql	não	sim
17	l	sim	2	sim	não	sim	sq	não	não
19	não	não	1	sim	não	não	ql	não	não
23	T + G	sim	3	não	não	não	qt	não	não
25	não	não	2	sim	sim	não	ql	não	não
26	não	não	2	sim	sim	não	qt	não	não
28	l	sim	3	sim	não	não	ql	não	não
31	l	sim	2	não	não	não	qt	não	não
32	l	não	1	sim	sim	não	sq	não	não
33	T	sim	2	não	não	não	sq	não	não
34	não	não	0	sim	não	não	ql	não	não
35	não	não	2	sim	não	não	ql	não	não
42	não	não	3	sim	não	não	ql	não	não
44	não	não	1	não	não	não	qt	não	não
45	T	sim	2	não	não	não	qt	não	não

ENEM 2009/2

Item	Gabarito	Competência	Habilidade	Habilidades Possíveis	Tema de Física Abordado (Geral)	Tema de Física Abordado (Específico)	Extensão Texto Base (linhas em 1/2 página)	Extensão Enunciado (linhas em 1/2 página)
5	E	1	3		2	2.2	16	2
14	E	2	5	5, 19 ou 23	5	5.9	8	3
17	E	5	18	18 ou 20	2	2.4	8	10
18	E	2	6	5 ou 6	5	5.9	3	11
19	D	5	17	5 ou 17	5	5.9	15	2
20	E	6	23		3	3.1	6	4
24	A	6	21	17 ou 21	7	7.11	10	2
27	D	6	20		2	2.13	8	2
30	D	6	20		5 e 7	5.9 e 7.4	8	5
31	E	1	2	1 ou 2	6	6.5	8	6
32	B	6	22	1 ou 22	6	6.5	9	3
35	A	6	21	17, 21 ou 22	7	7.4	15	8
37	D	6	22		6	6.4	14	3
38	D	6	21		7	7.6	9	6
39	B	6	21		7	7.9	11	2
45	B	2	5		5	5.10	8	4

Item	Linguagem não textual	interpretação de imagens / texto	Nível de Contextualização	Há cobrança explícita de Conteúdo	Necessita de memorização de fórmula	Laboratório é diferencial	quanti / quali/ semi	conexão cognitiva (conteúdos)	conexão cognitiva (áreas)
5	não	não	3	não	não	não	ql	não	sim
14	I	sim	2	sim	não	sim	ql	não	não
17	não	não	1	sim	sim	não	qt	não	não
18	T	sim	1	sim	não	não	ql	não	não
19	T + T + I	sim	3	não	não	não	qt	não	não
20	I	sim	3	não	não	não	ql	não	não
24	G	sim	1	sim	não	não	ql	não	não
27	I	não	1	sim	não	não	ql	não	não
30	não	não	3	sim	sim	não	qt	sim	não
31	não	não	1	sim	não	não	ql	não	não
32	I	sim	3	sim	não	não	ql	não	sim
35	I	sim	1	sim	sim	não	qt	não	sim
37	não	não	1	sim	não	sim	ql	não	sim
38	não	não	3	sim	sim	não	qt	não	não
39	I	não	3	sim	não	não	ql	não	não
45	I	sim	3	sim	não	sim	sq	não	não

ENEM 2010/1

Item	Gabarito	Competência	Habilidade	Habilidades Possíveis	Tema de Física Abordado (Geral)	Tema de Física Abordado (Específico)	Extensão Texto Base (linhas em 1/2 página)	Extensão Enunciado (linhas em 1/2 página)
47	A	1	1	1 ou 22	6	6.2	7	3
48	C	2	5	5 ou 17	5	5.9	13	3
50	A	1	3	3 ou 21	7	7.1	7	2
52	C	5	18	6, 18 ou 23	7	7.9	5	1
54	anulada	6	22	1 ou 22	6	6.2	12	2
56	A	6	23		3	3.1	9	3
58	D	6	21		7	7.7	16	2
59	D	6	23	21 ou 23	7	7.10	11	2
63	A	3	8	8, 19, 23 ou 26	3	3.1	7	3
68	E	2	6		5	5.14	9	2
70	A	5	17	5, 6, 17 ou 18	5	5.9	3	5
78	B	6	21		5	5.5	13	2
81	E	6	20		2	2.18	5	2
84	D	1	1	1 ou 18	6	6.2	10	3
89	D	6	23		3	3.1	8	3

Item	Linguagem não textual	interpretação de imagens / texto	Nível de Contextualização	Há cobrança explícita de Conteúdo	Necessita de memorização de fórmula	Laboratório é diferencial	quanti / quali/ semi	conexão cognitiva (conteúdos)	conexão cognitiva (áreas)
47	não	não	0	sim	não	não	ql	não	não
48	T	sim	3	sim	sim	não	qt	não	não
50	não	não	3	sim	não	não	ql	não	não
52	não	não	3	sim	sim	não	ql	não	não
54	I	sim	2	sim	não	não	ql	não	não
56	I	sim	3	não	não	não	ql	não	sim
58	não	não	3	sim	não	sim	ql	não	não
59	não	não	2	sim	não	sim	ql	não	não
63	I	sim	2	sim	não	não	ql	não	sim
68	I	sim	3	não	não	sim	qt	não	não
70	T	sim	3	sim	sim	não	qt	não	não
78	não	não	2	sim	não	sim	ql	não	não
81	não	não	3	sim	sim	sim	ql	não	não
84	não	não	1	sim	sim	não	sq	não	não
89	não	não	3	não	não	não	ql	não	não

ENEM 2010/2

Item	Gabarito	Competência	Habilidade	Habilidades Possíveis	Tema de Física Abordado (Geral)	Tema de Física Abordado (Específico)	Extensão Texto Base (linhas em 1/2 página)	Extensão Enunciado (linhas em 1/2 página)
48	A	6	23	17 ou 23	3	3.1	9	3
49	C	6	20		4	4.4	0	3
54	C	1	3		6	6.3	6	4
57	E	5	18	5 ou 18	5	5.5	25	3
58	E	1	1		6	6.5	7	2
59	E	2	5	5 ou 17	5	5.7	7	2
64	D	6	20	17 ou 20	2	2.3	16	6
65	C	5	18	18 ou 20	2	2.18	8	1
67	E	6	22	1, 18 ou 22	6	6.2	10	2
70	C	6	21		5	5.15	8	2
81	E	6	21	18 ou 21	5	5.15	6	2
84	B	2	5	5 ou 18	5	5.9	9	4
85	B	1	1	1 ou 22	6	6.5	8	3
86	E	6	23	18 ou 23	3	3.1	8	2
90	E	2	7	6,7,17 ou 18	2 e 7	2.18 e 7.6	13	1

Item	Linguagem não textual	interpretação de imagens / texto	Nível de Contextualização	Há cobrança explícita de Conteúdo	Necessita de memorização de fórmula	Laboratório é diferencial	quanti / quali/ semi	conexão cognitiva (conteúdos)	conexão cognitiva (áreas)
48	G	sim	2	sim	não	não	ql	não	não
49	I	sim	3	sim	não	sim	ql	sim	não
54	não	não	2	não	não	sim	ql	não	sim
57	não	não	1	sim	não	não	ql	não	não
58	não	não	2	sim	não	não	ql	não	não
59	T	sim	3	sim	sim	não	sq	não	não
64	G	sim	1	sim	sim	não	qt	não	sim
65	I + I	sim	3	sim	não	sim	ql	não	não
67	não	não	3	sim	não	sim	ql	não	não
70	não	não	3	sim	não	não	ql	não	sim
81	I	sim	3	sim	não	sim	ql	não	não
84	I	não	3	sim	sim	não	qt	não	não
85	não	não	3	sim	sim	sim	sq	não	não
86	não	não	3	sim	não	não	ql	não	não
90	I	sim	3	sim	não	sim	ql	sim	não

ENEM 2011

Item	Gabarito	Competência	Habilidade	Habilidades Possíveis	Tema de Física Abordado (Geral)	Tema de Física Abordado (Específico)	Extensão Texto Base (linhas em 1/2 página)	Extensão Enunciado (linhas em 1/2 página)
46	D	6	20	17 ou 20	1	1.4	12	2
54	C	5	17		1	1.5	19	2
56	C	6	21		5	5.15	11	4
60	A	2	6	5, 6 ou 17	5	5.9	12	5
63	E	6	22	1, 17 ou 22	6	6.2	12	1
66	C	6	23		7	7.9	8	2
67	E	5	17	17 ou 22	6	6.5	12	4
70	D	2	5		5	5.10	6	2
73	B	6	20		2	2.18	10	2
74	D	1	1	1, 18 ou 22	6	6.8	12	3
77	D	6	20	17, ou 20	2	2.3	17	2
78	B	2	7		1	1.5	10	1
80	D	6	23		3	3.1	14	3
84	A	1	1		6	6.5	7	2
86	C	6	23		3	3.2	3	3

Item	Linguagem não textual	interpretação de imagens / texto	Nível de Contextualização	Há cobrança explícita de Conteúdo	Necessita de memorização de fórmula	Laboratório é diferencial	quantitativo / qualitativo / semi	conexão cognitiva (conteúdos)	conexão cognitiva (áreas)
46	I	sim	3	não	não	não	ql	não	não
54	T	sim	1	não	não	não	qt	não	não
56	não	não	3	sim	não	sim	ql	não	não
60	T	sim	2	sim	sim	não	qt	não	não
63	I + I	sim	3	não	não	não	ql	não	não
66	não	não	3	sim	não	não	ql	não	não
67	G	sim	3	não	não	não	ql	não	não
70	I	sim	3	sim	não	sim	ql	não	não
73	I	sim	3	sim	sim	não	qt	não	não
74	não	não	3	sim	não	não	ql	não	não
77	T	sim	3	sim	sim	sim	sq	não	não
78	I	sim	3	não	não	não	ql	não	
80	não	não	2	sim	não	não	ql	não	sim
84	I	sim	3	sim	não	não	ql	não	não
86	I	sim	3	sim	não	não	ql	não	não

Explorando o Mundo das Avaliações de Larga Escala (Versão Preliminar)

Wanderley P. Gonçalves Jr e Marta F. Barroso

Cara colega, caro colega

Avaliar é uma das tarefas mais executadas por professores, de todos os níveis: no ensino fundamental, no ensino médio e no ensino superior. Na maior parte das vezes, essa tarefa é considerada pesada por todos nós.

Hoje, cada vez mais, também somos, de diferentes formas, avaliados externamente - ou melhor, nossos alunos o são. E os resultados são divulgados, na forma de rankings nos jornais, de premiações e outros.

Este trabalho que você tem em suas mãos foi elaborado pensando em compartilhar algumas das coisas que aprendemos, ao analisar avaliações em educação e ao pensar o que fazemos em sala de aula com os resultados dessas avaliações em mãos.

Esperamos que você chegue aonde chegamos: a pensar que a avaliação precisa ser objeto de reflexão por nós, professoras e professores, para que seus resultados possam ter consequências nos trabalhos que executamos em nosso dia a dia. E a principal consequência deveria ser nos permitir melhorar o nosso trabalho!

Esse texto propõe-se a discutir a avaliação da aprendizagem, e como analisamos esta avaliação. Começamos com uma breve revisão e comentários a respeito de processos e técnicas de medida. Inicia-se pela ideia de medição nas ciências físicas e a dificuldade de transpor esse conceito para as ciências sociais. Essa discussão é importante porque... exatamente o que queremos dizer com “avaliar a aprendizagem”?.

Tudo isso vem recheado com alguns exemplos da Física.

Em seguida, uma questão que sempre parece árida: lidar com as ideias numéricas e as técnicas envolvidas para pensar os resultados de avaliações. Em outras palavras, discutir um pouco sobre estatística: as ideias de média, de desvio padrão, de análise de dados, de fazer modelos para os dados com o “ajuste” de dados aos modelos. São apresentadas algumas formas de fazer essas discussões de forma quase automática, com auxílio de alguns softwares.

Finalmente, chega-se ao ponto: as técnicas de avaliação de aprendizagem! A medida de aprendizagem através da aplicação de testes (“provas”), e as teorias da psicometria para obter “notas”, ou escores, a partir desses testes, com a apresentação da Teoria Clássica dos Testes (TCT) e as ideias básicas da Teoria da Resposta ao Item (TRI), utilizada no ENEM.

Sumário

Fazer medidas	2
Exemplos de Medida: um exemplo de avaliação de aprendizagem em Física ..	5
Como fazer medidas: mais um exemplo da Física	7
Como apresentar resultados das medidas	9
Probabilidade: conceitos básicos	10
Melhor valor: medidas de tendência central	14
Medidas de dispersão	17
Função de distribuição, ou densidade de probabilidade	20
Distribuição de Gauss ou distribuição normal	21
Ajuste de dados - Regressão Linear	24
Testes: a Teoria Clássica de Testes	25
Testes: a Teoria da Resposta ao Item	27
Construir uma ICC (curva característica do item) e o modelo de Rasch	30
Modelos para a Teoria da Resposta ao Item	36

Fazer medidas

Ao trabalhar com ideias e grandezas em Física, ou em alguma das ciências ditas “exatas”, parece natural imaginar que tudo pode ser expresso em termos de grandezas mensuráveis que se relacionam por meio de equações.

As medidas são feitas, em geral, com dois objetivos em mente: descrever comportamentos (função descritiva), ou seja, estudar como os resultados experimentais se comportam, ou obter relações entre as variáveis medidas (função explicativa).

No campo das ciências naturais, e para ser mais específico, na Física (Alonso e Finn 1972)

“medir é um processo que nos permite atribuir um número a uma propriedade física como resultado de comparações entre quantidades semelhantes, sendo uma delas padronizada e adotada como unidade.”
(pág. 13)

Mas é muito comum imaginar que não se pode aplicar os raciocínios da física às demais ciências, como a psicologia, sociologia, economia, antropologia, ensino e aprendizagem em física, etc. Isso parece sugerir que as ciências como a Física constroem afirmações verdadeiras, enquanto as demais ciências não (Babbie 2005), como se houvesse um questionamento do status “científico” dessas outras áreas. Em outras palavras, será que o comportamento humano pode ser sujeito a medidas? Será que o “método científico” pode ser aplicado ao comportamento humano?

Vejamos: em qualquer área do conhecimento, as teorias existentes quase sempre resultam de uma combinação de processos dedutivos e indutivos, isso é, uma explicação inicial para uma observação é testada, reformulando a ideia inicial até a construção de um corpo teórico plausível e que explique as observações conhecidas. Neste processo, conceitos são utilizados, e muitas vezes para defini-los melhor é necessário propor uma forma de medi-lo, fazendo o que é chamado de definição operacional do conceito.

A generalização da ideia de medir pode ser feita (Allen e Yen 2002):

“A medida é a atribuição sistemática de números para indivíduos de um conjunto, tendo como objetivo a representação das propriedades desses indivíduos. Os números atribuídos a esses indivíduos devem ser obtidos através de procedimentos cuidadosamente prescritos e reproduzíveis. Por exemplo, testes de personalidade geram suas pontuações a partir do uso das mesmas instruções, perguntas, e procedimentos de pontuação para cada examinando. Essas pontuações não podem ser comparadas de forma significativa, se a cada um dos examinandos forem dadas instruções diferentes ou se diferentes procedimentos de pontuação forem utilizados para estabelecê-los. Na mensuração, os números são atribuídos de forma sistemática e podem possuir formatos variados. Por exemplo, pode-se atribuir-se às pessoas com cabelos vermelhos o número “1” e às pessoas com cabelos castanhos “2” como forma de medida. Neste caso, os números são atribuídos aos indivíduos de uma forma sistemática de maneira que as diferenças entre os escores representam as diferenças na propriedade que está sendo medida (cor do cabelo). Da mesma forma, dando-se a um examinando uma nota 98 na prova de matemática ou a nota 54 num teste de personalidade, estará se realizando uma medição, desde de que os números sejam sistematicamente atribuídos para representar as diferenças de desempenho nos testes de matemática ou de personalidade”.
(pág. 2)

Nas áreas da ciência em que o objeto de estudo é o conhecimento humano, é possível e muitas vezes desejável a realização de medidas sistemáticas. Por exemplo, as características sociais “cor de pele”, “idade”, “cidade natal”, “sexo”, “renda familiar” são facilmente mensuráveis.

Mas o que seria medir “aprendizagem em física”? Como seria medir a “religiosidade” de uma pessoa? Isto é, conceitos abstratos, atitudes, são passíveis de medida? Uma reflexão sobre isso é feita por Babbie (Babbie 2005)

“Deve-se reconhecer que todas estas medidas (todas medidas, aliás) são basicamente arbitrárias. O cientista social não pode descrever uma pessoa inequivocamente como “alienada” e outra como “não alienada”. Pessoas serão, ao invés, descritas como relativamente mais ou menos alienadas - ou seja, comparando uma com a outra. Esta característica, entretanto, não é prerrogativa das ciências sociais, como demonstram a “escala de dureza” usada nas ciências físicas, a “escala Richter” para terremotos, etc. Ninguém pode dizer que um metal é “duro” ou que um terremoto é “severo”, apenas que é mais “duro” ou mais “severo” que o outro.” (p. 59)

Muitas vezes, por motivos diversos, temos necessidade ou desejo de medir conceitos abstratos como classe social, racismo ou a aprendizagem em física.

Por serem abstratos, esses conceitos, em geral, não passam de ideias gerais na mente do pesquisador. Para que se possa medi-los, é necessário melhorar essas ideias, fornecer uma definição operacional, ou um conceito diferente. Começa-se levantando as características desses conceitos, reduzindo-se as características a indicadores empíricos específicos. Essa não será certamente a definição completa e indiscutível do conceito, mas representarão as características dele que julgamos úteis para a necessidade.

Por exemplo, ao se perguntar a professores de ensino médio o que entendem por aprendizagem em física, alguns afirmariam que saber física é saber resolver problemas numéricos de física; outros diriam que é solucionar problemas que necessitam de conhecimento dos conceitos físicos estudados, ou ainda, que a aprendizagem se revelaria se o aluno conseguisse aplicar os conceitos estudados em seu cotidiano. Cada uma dessas posições apresenta uma característica que pode ser utilizada para uma definição operacional do que seria aprendizagem em física.

Podemos olhar para definições acadêmicas mais conhecidas. Como exemplo, o PISA (Programa Internacional de Avaliação de Estudantes) avalia, em ciências, o quanto o aluno adquiriu de “letramento científico”, definido como “a capacidade de usar o conhecimento científico para identificar questões e chegar a conclusões baseadas em evidências para entender e ajudar a tomar decisões a respeito do mundo e as mudanças causadas a ele pela atividade humana”.

Já nos Parâmetros Curriculares Nacionais para o Ensino Médio, na área de Física, definem como aprendizagem em física “o resgate do espírito questionador do aluno e seu desejo de explorar o mundo, reconhecendo a física como cultura e como possibilidade de conhecer este mundo que o cerca”. Neste processo, o aluno deve desenvolver uma atitude reflexiva e autocrítica diante dos erros cometidos, gerenciar os conhecimentos adquiridos e compreender a predominância de aspectos técnicos e científicos na tomada de decisões sociais significativas e os conflitos gerados, nestes, pela negociação política.

O ENEM (Exame Nacional de Ensino Médio) entende que aprender física significa que o aluno adquiriu ou desenvolveu um conjunto de habilidades que, em conjunto, representam a aquisição e o desenvolvimento de algumas competências (aptidões) elencadas na Matriz de Referência desse exame.

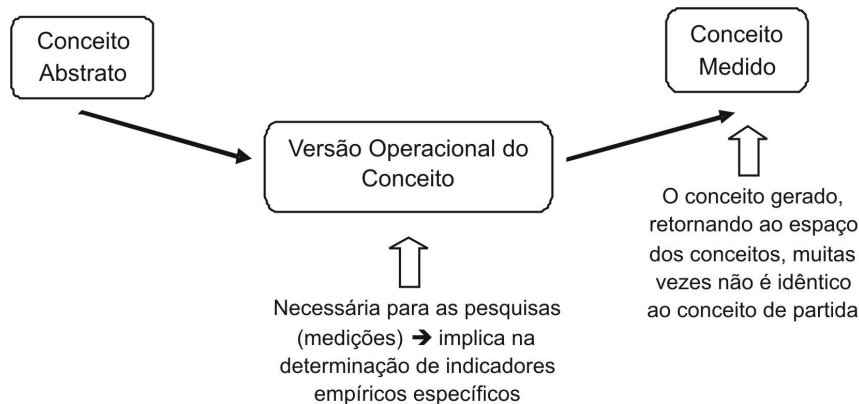
Em resumo, conceitos abstratos podem ser medidos se houver uma definição operacional plausível para eles elaborada pelas pessoas envolvidas; quase sempre não é possível construir-se um significado universal para esses conceitos. Por isso, não faz sentido dizer que o conceito operacional de um conceito abstrato está certo ou errado, mas sim dizer se ele é mais ou menos útil para responder às questões propostas.

É necessário refletir sobre as consequências desta discussão. A partir do momento que o pesquisador define o conceito operacional e quais aspectos desse conceito serão estudados, pode-se dizer que ele não coleta dados para o seu estudo, mas sim que cria esses dados.

Medida de tempo

O que é tempo? Essa pergunta é título de livros, de teses filosóficas, de discussões ardorosas... Mas existe alguma dúvida quanto ao fato que sabemos medir o tempo? Isto é, que temos uma definição operacional simples para ele? Contar o ângulo girado por um motor (relógio de pulso), as batidas de um metrônomo, entre outros!

A nossa conclusão então é que medir um conceito abstrato, muitas vezes, implica em “recriar” este conceito através de parâmetros concretos que sejam possíveis de ser medidos, gerando assim um novo conceito (o conceito operacional), que apesar de útil, não é a medida real do conceito abstrato em questão. De forma esquemática,



Define-se como a “operacionalização” de um conceito (Babbie 2005) o processo pelo qual são especificadas observações empíricas que podem ser tomadas como indicadores dos atributos contidos deste conceito.

Exemplos de Medida: um exemplo de avaliação de aprendizagem em Física

Pode-se agora exemplificar como um mesmo conceito pode ser tratado ou medido de diferentes formas, fornecendo diferentes visões do conceito.

No ensino médio, a discussão de movimentos de objetos próximos à superfície da Terra faz parte do conteúdo abordado na maior parte das escolas. Este tópico também é abordado no primeiro ano dos cursos universitários da área de ciência e tecnologia.

Tradicionalmente (a forma de trabalho da maior parte das escolas) este tópicos é ensinado com auxílio de exercícios que envolvem trabalho exaustivo com equações e gráficos na resolução de problemas numéricos. Coerentemente, a avaliação da aprendizagem desse conteúdo é feita através de provas e trabalhos com exercícios numéricos desse mesmo tipo.

Vejamos um exemplo.

Um corpo é lançado verticalmente para cima, em um local onde o efeito do atrito com o ar é desprezível. Se a velocidade de lançamento foi de 30 m/s qual a altura máxima atingida pelo corpo? (Considere $g = 10 \text{ m/s}^2$.)

O processo de resolução é tradicional e trivial. Utilizam-se as equações que descrevem movimentos com aceleração constante. Em particular, a chamada “equação de Torricelli”, $v^2 = v_0^2 + 2a\Delta s$, onde v é a componente da velocidade no eixo vertical, v_0 é a velocidade inicial, a é a aceleração e Δs o deslocamento realizado.

Sabendo que ao atingir a altura máxima a componente vertical da velocidade do corpo se anula, o procedimento habitual de resolução dos estudantes é:

$$v^2 = v_0^2 + 2a\Delta s$$

$$0^2 = 30^2 + 2(-10)\Delta s$$

$$20\Delta s = 900$$

$$\Delta s = 45\text{m}$$

Ao corrigir esse problema nos testes e provas, observa-se que o aluno, para o ponto de altura máxima e conseqüentemente de velocidade instantânea zero, substitui o valor da aceleração da gravidade por 10 m/s^2 .

Para um professor que utiliza este tipo de problema, medir a aprendizagem em física nesse conteúdo consiste na verificação da memorização e aplicação de equações em situações numéricas. Dentro dos parâmetros estabelecidos pelo professor para o conceito de aprendizagem deste conteúdo, há total coerência em suas ações e ele pode afirmar que o aluno que acerta o exercício aprendeu o conteúdo.

No entanto, se a concepção de aprendizagem dos conceitos físicos deste conteúdo implicar na exigência de uma compreensão mais conceitual das grandezas físicas envolvidas no fenômeno, pode-se propor outro tipo de exercício.

Observe as questões a seguir. Elas foram aplicadas a alunos calouros de um curso da área de ciências exatas da UFRJ, que na sua maioria são ingressam vindos de escolas tradicionais (comunicação privada, M.F.Barroso).

12) Uma bola é lançada verticalmente para cima. No ponto mais alto da trajetória da bola,

Resposta	Média	Total
sua aceleração é nula, e sua velocidade é não nula.	8.5%	8
sua velocidade e aceleração são nulas.	57.4%	54
sua velocidade é nula, mas a aceleração não é nula.	35.1%	33

Dos resultados, pode-se verificar que mais da metade dos estudantes responde que a aceleração no ponto mais alto da trajetória é nula. Em outras palavras, a repetição de exercícios numéricos durante o ensino médio não garantiu a compreensão

do conceito da aceleração da gravidade. E essa interpretação dos resultados é reforçada com a questão seguinte, aplicada ao mesmo grupo.

13) Considere duas situações:

Situação I - Uma bola é lançada para cima, verticalmente.

Situação II - Uma bola é largada do alto de uma torre.

Podemos afirmar que

Resposta	Média	Total
a aceleração depende da velocidade com que a bola é lançada na situação I, e da altura que é largada na situação II.	24.5%	23
na primeira situação, a aceleração é vertical e para cima, e na segunda a aceleração é vertical e para baixo.	51.1%	48
nada podemos afirmar sobre a aceleração, pois não temos nenhuma informação sobre as velocidades.	1.1%	1
nas duas situações, a bola tem a mesma aceleração.	11.7%	11
Nenhuma das respostas anteriores está correta.	12.8%	12

Esses resultados apontam para o fato que resolver problemas numéricos não garante a aprendizagem do conceito aceleração da gravidade. Em outras palavras, um professor que entende a compreensão do conceito como aprendizagem do mesmo não terá nos problemas numéricos bons parâmetros para sua avaliação.

Pesquisas em Ensino de Física indicam (Arons 1997) que existe uma dificuldade tremenda dos estudantes em situações como a contemplada nas duas questões respondidas pelos calouros da UFRJ. Logo, o conceito de aprendizagem em física adotado pelo professor exemplificado no primeiro problema (numérico) não é útil quando se deseja que o estudante adquira uma compreensão maior sobre o conceito de aceleração num da gravidade, pois, apesar de utilizar matematicamente o valor desta aceleração para o cálculo da altura atingida pelo corpo no primeiro problema apresentado, na maior parte das vezes ele não consegue relacionar a operação matemática com o conceito solicitado nas duas questões.

As duas formas de avaliar a aprendizagem, pela apresentação de um problema numérico e pela apresentação de questões conceituais, revelam resultados diferentes. Isto é, propusemos duas formas diferentes de operacionalizar a medida “aprendizagem de queda livre”. E obtemos resultados discrepantes. Em outras palavras, a operacionalização forneceu medidas diferentes, nenhuma certa ou errada, apenas mais ou menos útil para o objetivo do proponente da medida.

Como fazer medidas: mais um exemplo da Física

Vamos retornar aos exemplos de Física, para entender com exemplos mais simples a discussão de como tratar os dados.

Ao realizar uma medida para descrever um fenômeno - e vamos tomar como exemplo a determinação do período de um pêndulo simples (um pequeno objeto pendurado na extremidade de um fio) - é necessário entender que a medida a ser

realizada deve expressar, de forma clara, um resultado que seja compreensível (e muitas vezes reproduzível) por outra pessoa. Na prática, precisamos de uma unidade e, mais do que isso, precisamos apresentar um número e a exatidão deste número.

Sobre a “exatidão” de uma medida

Mesmo nas ciências exatas como a física, nem sempre conhecemos qual é o “valor verdadeiro” de uma medida. A medida que fazemos deve sempre se aproximar deste valor, mas não sabemos quão próxima deste valor está a medida feita. E para comparar resultados, avaliar teorias, é muito importante conhecer esta “incerteza” na medida - isso é, qual é a nossa avaliação da proximidade possível entre o valor verdadeiro e a medida que fazemos.

Uma maneira mais ou menos óbvia de estimar a exatidão de nossa medida é repeti-la, em iguais condições. É intuitivo que quanto mais vezes repetirmos a medida, melhor vai ser o resultado final (tratando os dados obtidos). Mas... a repetição de uma medida em idênticas condições não fornece resultados idênticos! Experimente medir algumas vezes com um cronômetro o período de oscilação de um pêndulo... Essas pequenas diferenças são flutuações estatísticas em nossos resultados - e as incertezas associadas são chamadas de erros aleatórios.

Há outro tipo de inexatidão da medida, que não podem ser corrigidos ou minimizados por repetição. O exemplo é um instrumento cuja calibração se altera com o tempo. Outro exemplo é você medir a altura da pessoa em diferentes momentos do dia (sabe-se que em geral ao longo do dia a pessoa “encolhe”). Esse tipo de inexatidão é denominada erro sistemático, e é difícil de ser evitado.

Os conceitos de precisão e acurácia estão associados a esses dois tipos de inexatidões experimentais. Veja o exemplo:

“Um jogador de futebol está treinando cobranças de penâlti. Ele chuta a bola 20 vezes, e 20 vezes acerta na trave do lado direito do goleiro. Ele é extremamente preciso, pois seus resultados não apresentam nenhuma variação em nenhuma das 20 vezes. Em compensação, sua acurácia é nula - ele nunca consegue acertar o “valor verdadeiro”, o gol.

Isso também tem efeitos em outros tipos de medida. Por exemplo, numa pesquisa eleitoral pode acontecer que o pesquisador entre numa sala com 20 pessoas que decidem informar ao pesquisador que votarão em um determinado candidato - o que não é necessariamente verdadeiro. Então, o levantamento de dados feito por este pesquisador é totalmente inacurado...

Em um laboratório de física, propusemos um experimento: cada um dos alunos mede um certo número de vezes o período de oscilação de um pêndulo simples com um cronômetro. Este cronômetro determinava o intervalo de tempo entre o ligar e desligar com medidas até centésimo de segundo.

Observou-se que era muito pouco provável a repetição dos valores encontrados na medida realizada por cada aluno. Essas medidas diferentes são resultado de muitos pequenos fatores, não controláveis pelo observador: a dificuldade de definir exatamente o final da oscilação, a demora ou a rapidez em apertar o botão de ligar e desligar, entre outros. E aí, o que devemos fazer? Qual das medidas é o período do pêndulo?

Como apresentar resultados de medidas

Para se determinar a cor preferida de um grupo de 30 pessoas, pode-se fazer a pergunta a cada uma delas. Para pesquisar a intenção de voto para Presidente da República, é praticamente impossível consultar todos os eleitores (a *população* ou o *universo estatístico*).

*população
amostra*

Recorre-se neste caso ao que se denomina de *amostra*, isto é, um grupo de indivíduos ou objetos pertencente ao universo pesquisado; a pergunta é feita a este subgrupo, e espera-se chegar ao resultado que reflita o todo. Cada pessoa da amostra é denominado *indivíduo* ou *objeto*. No caso da intenção de voto, as pessoas são os indivíduos.

Imagine que uma construtora pretenda lançar um empreendimento imobiliário num bairro da cidade. Para isso, a empresa faz uma pesquisa para sondar a preferência dos possíveis compradores em relação ao tamanho dos apartamentos, número de vagas de garagem, cor de fachada, área de lazer, sistema de segurança, etc. Cada uma dessas características é uma variável da pesquisa.

Na variável número de vagas de garagem, as opções podem ser “nenhuma”, “uma”, “duas” ou “mais de duas” vagas. Diz-se que esses são os *valores* ou *realizações da variável* “numero de vagas de garagem”.

Essas variáveis, por sua vez, podem ser classificadas em dois grupos.

*variáveis qualitativas
ordinais e nominais*

O primeiro deles, das variáveis qualitativas, é composto pelas variáveis que apresentam como possíveis valores uma qualidade ou atributo dos indivíduos; como exemplo, a cor da pele, o esporte favorito, o grau de instrução, o gênero. Se essas variáveis qualitativas possuem uma ordem em seus valores, como ocorre com grau de instrução (fundamental, médio, superior), essa é uma variável qualitativa *ordinal*. No caso do esporte favorito, essa variável é uma variável qualitativa *nominal*.

*variáveis quantitativas
discretas e contínuas*

No segundo grupo estão as variáveis quantitativas, que possuem valores numéricos como possíveis realizações. Os exemplos são a idade do indivíduo, sua altura, seu peso, o número de irmãos ou dependentes. Essas variáveis também podem ser divididas em dois subgrupos: o das variáveis quantitativas *discretas*, que constituem as variáveis numéricas representadas por números inteiros como o número de irmãos ou dependentes, e o das variáveis quantitativas *contínuas*, quando a medida pode ser representada por números reais, como a altura ou o peso do indivíduo.

*frequência
absoluta e relativa*

Após a realização de medidas, costuma-se apresentar o resultado de diversas formas. Uma das formas corresponde à apresentação do número de vezes que um valor da variável é obtido, a denominada *frequência absoluta* do valor. Quando se registra a frequência absoluta de um valor em relação ao total de valores obtidos, tem-se a *frequência relativa*.

Por exemplo, suponha que em uma pequena sala de aula tenha sido feito um levantamento sobre a descendência étnica por parte de mãe de 10 alunos e que o resultado seja o seguinte: três alunos descendentes de mãe negra, um de japonesa, quatro de italiana e dois de portuguesa. A frequência absoluta da variável “descendência” de cada um dos seus valores é: negros, 3; japoneses, 1; italianos, 4; portugueses, 2. Já a frequência relativa da descendência italiana é: 4 em 10 ou 0,4 ou 4/10 ou 2/5 ou 40% (os dados estão apresentados na Tabela 1).

Tabela 1. Exemplo de frequência de uma medida qualitativa nominal.

Descendência étnica por parte de mãe			
	frequência	frequência relativa	frequência percentual
afrodescendente	3	0,3 (3/10)	30%
japonesa	1	0,1 (1/10)	10%
italiana	4	0,4 (4/10)	40%
portuguesa	2	0,2 (2/10)	20%
Total	10	1 (10/10)	100%

Probabilidade: conceitos básicos

Ao lançar uma moeda para o alto, sabe-se intuitivamente que a “chance” de cada uma das suas faces cair para cima é de 1 para 2, ou de 50%. Se rolarmos um dado, a “chance” de o número dois ficar virado para cima é de uma em seis ou 1/6. O termo “chance”, entre aspas, refere-se ao que se denomina de *probabilidade*. Afirmar que a probabilidade de se obter “cara”, no lançamento da moeda é de 50% significa dizer que se ela for lançada um grande número de vezes, a quantidade de vezes que se obterá a face com “cara” voltada para cima será aproximadamente a metade dos lançamentos. Da mesma forma, ao se jogar o dado, também um grande número de vezes, ter-se-á que o número dois ficará na face superior do dado cerca de 1/6 das vezes. À medida que o número de eventos realizados para um determinado fenômeno aumenta, aumenta também a proximidade entre os resultados obtidos estatística e experimentalmente. De acordo com Young (Young 1962):

“No problema do lançamento da moeda, deve-se ressaltar que a razão entre o número de vezes que se obtém cara e o número total de eventos se aproxima de 0,5 à medida que o número de eventos se torna muito grande. Isso não é a mesma coisa que dizer que o número de caras obtidas se aproxima do número de coroas. Por exemplo, em 100 lançamentos, um resultado razoável a ser obtido seria obter cara 52 duas vezes. Para 10.000 lançamentos um provável resultado seria obter 5020 caras. Nesse segundo caso, a razão se aproxima muito mais de 0,5 do que no primeiro caso, mas a diferença entre o número de caras e coroas obtidas é maior. Na realidade, pode-se mostrar que a diferença entre o número de caras e o número de coroas obtidas tende a se tornar cada vez maior, independente do fato que a razão de cada uma delas em relação ao total de eventos se aproxima de 0,5. Portanto, se você está tirando cara e coroa com alguém e está perdendo, você não necessariamente recuperará suas perdas após um grande número de lançamentos. Existe 50% de chance de você perder cada vez mais e mais.” (p. 24)

distribuição de probabilidades

Ao jogar 10 moedas para o alto ao mesmo tempo, pode-se contar facilmente o número de caras e coroas obtidas. Mas se o que se deseja saber é a probabilidade de se obter “n” caras e “(10-n)” coroas nessa brincadeira, sendo “n” um número inteiro entre 0 e 10, a resposta corresponde a um conjunto de números para cada valor de “n”. Estes

números podem ser pensados como uma função de “n”, representada por $f(n)$ e denominada distribuição de probabilidade. No caso das 10 jogadas das moedas, se obtivemos 4 caras e 6 coroas, escrevemos:

$$f(\text{cara}) = 0,4$$

$$f(\text{coroa}) = 0,6$$

Em outras palavras, distribuição de probabilidade pode ser definida como uma representação do conjunto de probabilidades de todos os eventos associados a um espaço amostral.

Como a distribuição de probabilidade é definida para um intervalo determinado de valores de “n”, a soma de todas as probabilidades de todos os valores deve ser igual a 1, ou seja

$$\sum_n f(n) = 1$$

Vamos considerar uma turma de estudantes (nomes fictícios), como exemplo de distribuição de probabilidades discreta. Na Tabela 2, apresentamos a lista dos alunos e suas respectivas notas numa prova cujo grau variava de 0 a 10, com intervalos de 0,5 pontos.

Tabela 2. Notas dos alunos (máximo possível 10 pontos)

Número	Nome do aluno	Grau obtido
1	Agnes Paula	3,0
2	Ana Maria	2,0
3	Ana Paula	3,0
4	Arthur Ananias	1,5
5	Débora Mortícia	8,0
6	Felipe Barbosa	8,0
7	Fernanda Lima	6,0
8	Glenda Rodrigues	5,5
9	Helena Laura	7,0
10	João Pedro	9,5
11	Laura Miller	2,0
12	Luiz Felipe	8,0
13	Luiza Helena	5,5
14	Marina Lima	9,5
15	Mauro Rolando	6,0
16	Nilo Egípcio	4,5
17	Roberta Carla	6,0
18	Tatiana Alegre	1,5
19	Vanessa Guimarães	7,5
20	Vitor Peçanha	4,5

Para determinar a distribuição de probabilidades dos dados apresentados na Tabela 2, colocamos as notas na primeira coluna e na segunda coluna o número total de alunos que tiraram aquela nota. Contamos da Tabela 2 quantos alunos tiraram cada nota para preencher a segunda coluna da Tabela 3.

Tabela 3. Número de alunos por nota.

Nota	Número de alunos
1,5	2
2,0	2
3,0	2
4,5	2
5,5	2
6,0	3
7,0	1
7,5	1
8,0	3
9,5	2
Total	20

Na Figura 1, mostra-se de forma gráfica a informação contida na Tabela 3. No eixo horizontal, tem-se o valor da nota. Na vertical, o número de alunos que obtiveram aquela nota. Os dados são apresentados na forma de barras - e esse gráfico é chamado de histograma das notas desta turma. Observe que para o valor 8 o número é 3.

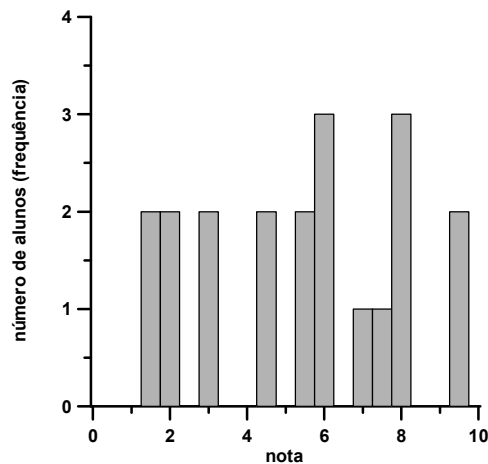


Figura 1. O histograma das notas dos alunos (da Tabela 3).

A distribuição de probabilidade $f(n)$ é obtida tomando-se a razão entre o número de alunos com uma determinada nota e o número total de alunos, no caso 20. Então para a nota 1,5, temos que a distribuição de probabilidade será $f(1,5) = 2/20 = 1/10$, para a nota 6,0 a distribuição de probabilidade será $f(6,0) = 3/20$. Denominando-se, então, por n cada nota obtida no conjunto de provas considerado e por $f(n)$ a distribuição de probabilidade para cada valor de n , obtém-se a Tabela 4 com a distribuição de probabilidades.

A soma de todos os valores de $f(n)$ deve ser 1. Dos dados da Tabela 4:

$$\begin{aligned}
 f(0,5) + f(2,0) + f(3,0) + f(4,5) + f(5,5) + f(6,0) + f(7,0) + f(7,5) + f(8,0) + f(9,5) &= \\
 = 0,1 + 0,1 + 0,1 + 0,1 + 0,1 + 0,15 + 0,05 + 0,05 + 0,15 + 0,10 &= 1
 \end{aligned}$$

Tabela 4. Distribuição de probabilidades das notas.

n (Nota)	f(n)
1,5	0,10
2,0	0,10
3,0	0,10
4,5	0,10
5,5	0,10
6,0	0,15
7,0	0,05
7,5	0,05
8,0	0,15
9,5	0,10
Total	1,00

Pode-se representar a distribuição de probabilidades da Tabela 4 por um histograma de frequências relativas, como na Figura 2.

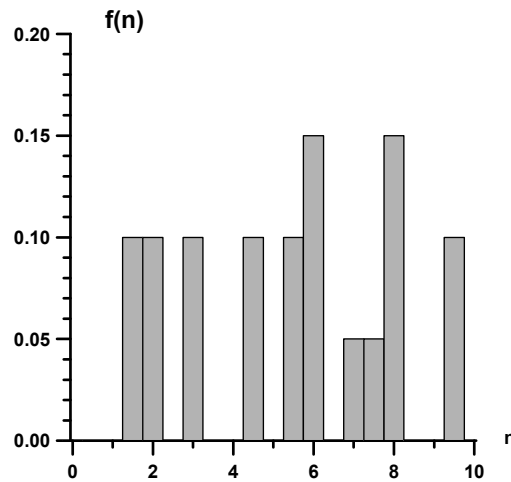


Figura 2. Distribuição de probabilidade de notas dos alunos

Vejamos outro exemplo. Um aluno, no laboratório, faz 10 medidas do período de oscilação de um pêndulo, e os valores por ele obtidos são anotados na Tabela 5.

Tabela 5. Medidas do período de oscilação de um pêndulo.

nº da medida	período (s)
1	3,19
2	3,31
3	3,26
4	3,33
5	3,32
6	3,28
7	3,41
8	3,29
9	3,34
10	3,42

Esses dados podem assumir qualquer valor dentro das limitações do cronômetro (de centésimo de segundo).

Como cada uma das dez medidas obtidas é anotada na tabela uma única vez, a probabilidade de obtenção de cada uma das medidas é igual e corresponde ao valor $1/10$, como mostrado na Tabela 6.

Tabela 6. Distribuição de probabilidade dos períodos obtidos.

período (s)	probabilidade
3,19	$1/10=0,1$
3,26	$1/10=0,1$
3,28	$1/10=0,1$
3,29	$1/10=0,1$
3,31	$2/10=0,2$
3,32	$1/10=0,1$
3,33	$1/10=0,1$
3,34	$1/10=0,1$
3,41	$1/10=0,1$
3,42	$1/10=0,1$

Novamente, a soma de todos os valores dá 1.

Melhor valor: medidas de tendência central

Quando queremos estudar ou apresentar as características de uma medida cujo resultado varia (a nota do aluno na prova, o período do pêndulo em medidas sucessivas), é interessante conseguir apresentar a informação com um pouco mais de simplicidade do que com uma tabela ou um gráfico. Por exemplo, podemos apresentar a idade de um grupo de alunos informando a média dessas idades, como se fosse uma idade representativa do grupo.

Na medida do período do pêndulo, um bom valor para responder à pergunta “qual é o valor do período deste pêndulo?” seria fornecer o valor médio das medidas realizadas. Em outras palavras, resumir um conjunto de dados a um valor representativo, que representa a “tendência central” do grupo de medidas.

medidas de tendência central

Algumas das medidas de tendência central são a *média aritmética* ou abreviadamente *média* (a mais conhecida), a *mediana*, a *moda* e a *média harmônica* ou média ponderada. Cada uma dessas medidas adequa-se a uma situação, com vantagens e desvantagens.

Imaginemos um conjunto de N dados

$$\{x_1, x_2, x_3, x_4, \dots, x_N\}$$

A média aritmética é obtida somando-se todos os valores e dividindo pelo total de valores; ou seja,

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

média aritmética

Por exemplo, suponha que numa sala de aula de jovens e adultos exista um grupo de 10 alunos com 24, 27, 31, 32, 33, 37, 39, 40, 45 e 47 anos. A média aritmética das idades desse grupo pode ser calculada por:

$$\bar{I} = \frac{24 + 27 + 31 + 32 + 33 + 37 + 39 + 40 + 45 + 47}{10} \Rightarrow \bar{I} = \frac{355}{10} = 35,5$$

Se os valores $X_1, X_2, X_3, \dots, X_k$ ocorrem com as frequências $f_1, f_2, f_3, \dots, f_k$, a média aritmética, neste caso, pode ser obtida substituindo os f_1 valores X_1 por $f_1 \cdot X_1$ e assim sucessivamente:

$$\bar{X} = \frac{f_1 \cdot X_1 + f_2 \cdot X_2 + f_3 \cdot X_3 + \dots + f_k \cdot X_k}{f_1 + f_2 + f_3 + \dots + f_k} = \frac{\sum_{i=1}^k f_i \cdot X_i}{\sum_{i=1}^k f_i}$$

onde n é a soma dos f_i , a frequência total ou o total de número de eventos.

Como exemplo, tomando-se as notas obtidas pelos alunos e suas respectivas frequências na Tabela 3, pode-se calcular a média aritmética como se segue:

$$\bar{N} = \frac{2 \cdot 1,5 + 2 \cdot 2,0 + 2 \cdot 3,0 + 2 \cdot 4,5 + 2 \cdot 5,5 + 3 \cdot 6,0 + 1 \cdot 7,0 + 1 \cdot 7,5 + 3 \cdot 8,0 + 2 \cdot 9,5}{2 + 2 + 2 + 2 + 2 + 3 + 1 + 1 + 3 + 2} =$$

$$\bar{N} = \frac{3,0 + 4,0 + 6,0 + 9,0 + 11 + 18 + 7,0 + 7,5 + 24 + 19}{20}$$

$$\bar{N} = \frac{108,5}{20} = 5,3$$

Sendo uma medida de tendência central, o cálculo da média aritmética busca através de um único número apresentar as características de um determinado grupo de números. Porém, em algumas situações, a presença de um valor muito maior ou muito menor que o restante do grupo pode fazer com que ela não consiga traçar o perfil correto desse grupo.

Por exemplo, se num grupo de seis pessoas que fizeram uma prova valendo 100 pontos, as notas obtidas foram respectivamente 2, 3, 1, 1, 2 e 50, a média aritmética desses valores será 9,8:

$$\bar{N} = \frac{2 + 3 + 1 + 1 + 2 + 50}{6} = \frac{59}{6} = 9,8$$

Observe que o valor obtido não representa bem as características desse grupo em termos de nota. Então, neste caso, a média não é uma medida de tendência central apropriada. Provavelmente a forma de comunicar o resultado seria a média eliminando o valor espúrio (50) e o valor separadamente: o grupo tem duas partes, um aluno tirou 50 em 100 e o resto teve média 1,8 em 100.

Em alguns casos deseja-se calcular a média de um conjunto de valores em alguns desses valores são mais importantes que outros, e essa importância é representada por um “peso”. Para se realizar, então, essa operação levando-se em conta essas “diferentes importâncias” define-se a média aritmética ponderada.

Consideremos os números $X_1, X_2, X_3, \dots, X_N$ associados respectivamente aos pesos $P_1, P_2, P_3, \dots, P_N$. A média aritmética ponderada desses valores é dada por:

$$\bar{X} = \frac{P_1 \cdot X_1 + P_2 \cdot X_2 + P_3 \cdot X_3 + \dots + P_n \cdot X_n}{P_1 + P_2 + P_3 + \dots + P_n} = \frac{\sum_{i=1}^n P_i \cdot X_i}{\sum_{i=1}^n P_i}$$

média ponderada

Como exemplo, pense em um professor que tenha aplicado duas provas e uma atividade de laboratório durante um bimestre. À primeira prova, objetiva, ele atribuiu peso 1. À segunda, peso 2, por ser discursiva; e atribuiu peso 3 ao desempenho na atividade de laboratório que, além de consumir maior tempo, necessitava um maior grau de atenção e criatividade. Gabriela, uma de suas alunas, obteve para essas avaliações, respectivamente, as notas 6,0, 4,5 e 7,0. Calculando-se a média aritmética ponderada das notas desta aluna obtém-se:

$$\bar{N} = \frac{6,0 \cdot 1 + 4,5 \cdot 2 + 7,0 \cdot 3}{1 + 2 + 3} = \frac{6,0 + 9,0 + 21}{6} = 6,0$$

Antes de prosseguir, é útil observar-se algumas das propriedades das médias aritméticas.

a) A soma algébrica dos desvios de um conjunto de números X_i , em relação à média aritmética, é zero:

$$\sum_{i=1}^n (X_i - \bar{X}) = 0$$

b) A soma dos quadrados dos desvios de um conjunto de X_i , em relação a um número qualquer “a”, é um mínimo se e somente se $a = \bar{X}$:

$$\sum_{i=1}^n (X_i - a)^2 \geq 0 \quad \text{é mínimo se e somente se } a = \bar{X}.$$

c) Se f_1 números têm média m_1 , f_2 números têm média m_2, \dots, f_k números têm média m_k , a média de todos os números é dada por

$$\bar{X} = \frac{f_1 \cdot m_1 + f_2 \cdot m_2 + f_3 \cdot m_3 + \dots + f_k \cdot m_k}{f_1 + f_2 + f_3 + \dots + f_k} = \frac{\sum_{i=1}^k f_i \cdot m_i}{\sum_{i=1}^k f_i}$$

Outro valor de tendência central é a mediana de um conjunto de números: se esses números são organizados em ordem de grandeza, é o valor que ocupa a posição central dos números. No caso desse conjunto de números possuir um número de elementos ímpar, a mediana ocupa o valor central dessa distribuição ordenada. No entanto, se o número de elementos for par, a mediana é obtida através da média aritmética dos dois valores centrais.

mediana

Por exemplo, um aluno durante um bimestre obteve as seguintes notas nas atividades avaliativas de uma determinada disciplina: 10, 9, 8, 7 e 10. Para obter-se a mediana desse conjunto de números necessita-se, inicialmente, ordenar esses valores, ou seja: 7, 8, 9, 10, 10.

Com a ordenação feita, a mediana será o valor que ocupa a posição central dos números, ou seja, o número 9.

No entanto, ao se procurar a mediana das notas obtidas pelos alunos na Tabela 3.1, não teremos um valor em posição central. Logo, para determinação da mediana deve-se fazer a média aritmética dos dois valores centrais, ou seja:

1,5 1,5 2,0 2,0 3,0 3,0 4,5 4,5 5,5 5,5 6,0 6,0 6,0 6,0 7,0 7,5 8,0 8,0 8,0 8,0 9,5 9,5

$$\tilde{X} = \frac{5,5 + 6,0}{2} = 5,8$$

Geometricamente, a mediana é o valor de “X” (abscissa) correspondente à vertical que divide o histograma (dos dados X_i) em duas partes de áreas iguais.

Chama-se moda de um conjunto de dados o valor que ocorre com maior frequência. A moda pode não existir e, mesmo que exista, pode não ser única. E esse é a última medida de tendência central usual.

moda

Quando uma distribuição de probabilidade possui somente uma moda ela é denominada de unimodal.

Medidas de dispersão

O conhecimento do valor médio da grandeza medida não é suficiente para descrever toda a informação obtida com a medida. Com apenas esta informação apenas, não se é capaz de discernir que tipo de medida foi feita, o que acontecerá se ela for repetida, não será possível dizer se a distribuição de probabilidades é simétrica, se é muito “larga” (isto é, se a faixa de valores medidos é grande), etc.

Quando se possui os valores das medidas X_i e sua respectiva média \bar{X} , pode-se saber o quão distante cada uma das medidas encontra-se da média, isto é, o quanto os dados numéricos tendem a se dispersar em torno do valor médio. A descrição quantitativa desse distanciamento, denominado de dispersão, fornece uma idéia da precisão dessas medidas.

Por exemplo, suponha que um professor é contratado para uma substituição num curso livre em que as aulas devem conter atividades recreativas. Como ele não conhece a escola, informam-lhe que a turma é constituída de um grupo de 6 pessoas cuja média de idade é de 20 anos. A informação da média de idades, no entanto, não é suficiente, pois se pode ter grupos com essa média de idade com características completamente diferentes:

Grupo 1: 25 anos; 25 anos; 25 anos; 25 anos; 25 anos; 25 anos.

média aritmética das idades: $\bar{I}_1 = \frac{25+25+25+25+25+25}{6} = \frac{150}{6} = 25 \text{ anos}$

Grupo 2: 28 anos; 25 anos; 27 anos; 25 anos; 24 anos; 21 anos.

média aritmética das idades: $\bar{I}_2 = \frac{28+25+27+25+24+21}{6} = \frac{150}{6} = 25 \text{ anos}$

Grupo 3: 10 anos; 64 anos; 51 anos; 14 anos; 8 anos; 1 ano.

média aritmética das idades: $\bar{I}_3 = \frac{10+64+51+14+8+1}{6} = \frac{150}{6} = 25 \text{ anos}$

Pode-se observar que a medida de tendência central (média aritmética) utilizada no exemplo não é suficiente para caracterizar bem os grupos, em particular o terceiro. Neste caso faz-se necessária a utilização de medidas que expressem o grau de dispersão de um conjunto de dados. Dessas medidas de dispersão, as mais utilizadas são a variância e o desvio padrão.

Não adianta pensar em usar a soma dos desvios, pois essa soma dá zero sempre (veja o comentário ao final da discussão sobre médias, e faça o cálculo nos exemplos acima!)

A idéia fundamental da variância (σ^2) é se tomar os desvios dos valores X_i em relação à média aritmética desses valores. A soma dos desvios simples, $d_i = X_i - \bar{X}$, é nula; poder-se-ia pensar em somar os valores absolutos desses desvios, ou então o quadrado desses desvios. Por motivos de facilidade de cálculo, considera-se a variância como sendo a média dos quadrados dos desvios, ou seja:

variância

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

Então, calculando-se a variância para os grupos 1, 2 e 3 do exemplo anterior, tem-se:

Grupo 1: 25 anos; 25 anos; 25 anos; 25 anos; 25 anos; 25 anos.

$$\bar{I}_1 = 25 \text{ anos}$$

$$\sigma^2 = \frac{(25-25)^2 + (25-25)^2 + (25-25)^2 + (25-25)^2 + (25-25)^2 + (25-25)^2}{6} = 0$$

Observe que pelo fato de todos os valores serem iguais, a variância é 0. Este resultado permite afirmar que não há dispersão, que o grupo é completamente homogêneo no que se refere à idade.

Grupo 2: 28 anos; 25 anos; 27 anos; 25 anos; 24 anos; 21 anos.

$$\bar{I}_2 = 25 \text{ anos}$$

$$\sigma^2 = \frac{(25-28)^2 + (25-25)^2 + (25-27)^2 + (25-25)^2 + (25-24)^2 + (25-21)^2}{6}$$

$$\sigma^2 = \frac{(-3)^2 + (0)^2 + (-2)^2 + (0)^2 + (1)^2 + (4)^2}{6}$$

$$\sigma^2 = \frac{9+0+4+0+1+16}{6} = \frac{30}{6} = 5$$

Grupo 3: 10 anos; 64 anos; 51 anos; 14 anos; 8 anos; 1 ano.

$$\bar{I}_2 = 25 \text{ anos}$$

$$\sigma^2 = \frac{(25-10)^2 + (25-64)^2 + (25-51)^2 + (25-14)^2 + (25-8)^2 + (25-1)^2}{6}$$

$$\sigma^2 = \frac{(15)^2 + (-39)^2 + (-26)^2 + (11)^2 + (17)^2 + (24)^2}{6}$$

$$\sigma^2 = \frac{225+1521+676+121+289+576}{6} = \frac{3408}{6} = 568$$

Com os valores obtidos da variância para cada grupo é possível diferenciar a dispersão de cada um deles. Observa-se que o grupo 1 não possui dispersão, enquanto o grupo 3 apresenta uma dispersão muito maior que o 2. Costuma-se dizer que o grupo 2 é mais homogêneo que o grupo 3 ou que o 3 é mais heterogêneo que o 2.

No entanto, pelo fato da variância ser calculada com os desvios $d_i = X_i - \bar{X}$ elevados ao quadrado, não é possível expressar a variância na mesma unidade dos valores da variável. Então, para que isso seja possível, define-se a medida de dispersão chamada de desvio padrão.

desvio padrão

O desvio padrão é a raiz quadrada da variância. A sua utilização facilita a interpretação dos dados por se expresso na mesma unidade dos valores observados.

Matematicamente tem-se:

$$\sigma = \sqrt{\sigma^2}$$

Então, para os grupos 1, 2 e 3, tem-se:

Grupo 1: $\sigma_1 = \sqrt{0} = 0$ ano

Grupo 2: $\sigma_2 = \sqrt{5} = 2,2$ anos

Grupo 3: $\sigma_3 = \sqrt{568} = 23$ anos

Em relação a essas medidas de dispersão, observa-se que:

- a) A variância e o desvio padrão são sempre positivos.
- b) Quando todos os valores da variável são iguais, o desvio padrão é 0.
- c) Quanto mais tende a zero o desvio padrão, mais homogênea é a distribuição de valores da variável.
- d) O desvio padrão é expresso na mesma unidade da variável.

Função de distribuição, ou densidade de probabilidade

Um conceito importante na apresentação de dados e resultados de medidas é a forma como são apresentados os resultados. Já vimos que é muito mais ilustrativo apresentar um histograma das notas dos alunos de uma turma do que apresentar a tabela dessas notas. Uma extensão da ideia de apresentação de resultados no formato de histogramas é a apresentação de “funções de distribuição”, ou densidade de probabilidade, dos valores ou dados.

Imaginemos que estamos tratando de todas as turmas de um grande colégio, com 1000 alunos. As médias finais desses alunos podem ser apresentadas em intervalos discretos, de 5 em 5 pontos (0; 5; 10; 15; ...), ou de forma quase contínua, de 0,1 em 0,1 pontos (0,1; 0,2; 0,3; ...; 99,9; 100) entre os valores de 0 a 100. Ao fazermos o histograma neste segundo caso, é conveniente agrupar os valores em intervalos maiores; nesse caso, há uma acumulação de dados em cada uma das “barras” que compõem o histograma. Por exemplo, com notas entre 50,0 e 50,4 temos as notas da Tabela 7.

Tabela 7. Médias finais de uma faixa numa escola com 1000 alunos

Nota	Número de alunos
∴ (de 0 a 49,9)	(253)
50,0	10
50,1	15
50,2	8
50,3	7
50,4	20
∴ (até 100)	(687)

Na “barra” do histograma que contém os dados entre 50,0 e 50,4 devem ser representados 60 alunos ($60=10+15+8+7+20$). Em outras palavras, o gráfico deve ser lido

multiplicando-se o tamanho do intervalo de agrupamento com o valor da abscissa (para a interpretação do valor total de alunos).

Distribuição de Gauss ou distribuição normal

Os histogramas de resultados de dados experimentais que são quase contínuos, como o caso das médias da grande escola ou o caso de fazermos 400 medidas de um período de um pêndulo com um cronômetro digital, transformam-se em diagramas de barras que em sua extremidade superior assumem o aspecto de uma curva contínua. Essa curva recebe o nome de distribuição, representando uma distribuição de probabilidades de obter o resultado na faixa considerada. Algumas formas que esta curva assume se repetem muito; em particular, em quase todos os casos que em os erros são instrumentais (aleatórios ou ao acaso), esta curva pode ser descrita por uma função particular, a função gaussiana. A distribuição cuja curva é uma função gaussiana é denominada distribuição de Gauss, ou normal.

A importância desse tipo de distribuição deve-se

- ao fato dela descrever a distribuição dos erros aleatórios em muitos tipos de medidas em situações físicas, biológicas e sociais;
- à possibilidade de se demonstrar que, mesmo que os erros individuais não sigam esta distribuição, as médias desses grupos de erros serão distribuídas de uma forma que se aproximem da distribuição de Gauss quando se trata de grupos com um grande número de elementos;
- ao fato de ser fundamental para a inferência estatística.

Numa descrição histórica sobre a construção da curva normal, um professor conhecido que trabalha em avaliações educacionais, o prof. Luiz Pasquali, descreve:

“A lei dos grandes números de Bernoulli diz o seguinte: numa situação de eventos casualóides, onde as alternativas são independentes, obter coroa em lances de uma moeda de cara e coroa, tem a probabilidade matemática exata de 50% (porque são somente dois eventos possíveis: cara ou coroa), mas na prática esta probabilidade de 50% é apenas aproximada. E essa aproximação é tanto mais exata quanto maior forem as tentativas que você fizer de lançar a moeda, chegando a quase atingir os exatos 50% se você lançar a moeda infinitas vezes. Isto é, quanto mais lances você fizer, menor será o desvio em relação à média de 50% que o resultado irá produzir. Isso quer dizer que os erros (desvios) serão menores e menores na medida em que sobe o número de lances. Desvios grandes são raros e desvios pequenos frequentes; quanto menores os desvios, mais frequentes eles serão, de sorte que, aumentando as tentativas (os lances), aumenta o número de desvios pequenos, sobrepujando cada vez mais os desvios grandes, de tal sorte que, no limite, haverá quase somente desvios pequenos, sendo o desvio 0 o menor deles e, por consequência, o mais frequente.

Moiivre assumiu essa idéia de Bernoulli e disse: erros grandes são mais raros que erros pequenos. Assim, quanto menores os erros, mais freqüentes eles serão e quanto maiores, menos freqüentes. Dessa forma, os erros se distribuem eqüitativamente em torno de um ponto modal, a média, formando uma curva simétrica com pico na média e caindo rapidamente para as caudas à esquerda (erros que subestimam a média) e à direita (erros que superestimam a média).” (Luiz Pasquali, Psicometria, 1996, p. 71 e 72).

A distribuição a que se refere a citação é justamente a distribuição gaussiana ou normal. Ela é uma curva simétrica e contínua, na forma de um “sino”, que possui a seguinte expressão matemática:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Esta função possui dois parâmetros: \bar{x} , que é o valor médio e σ^2 , que é sua variância.

Plotando-se a curva da função P(x), tem-se:

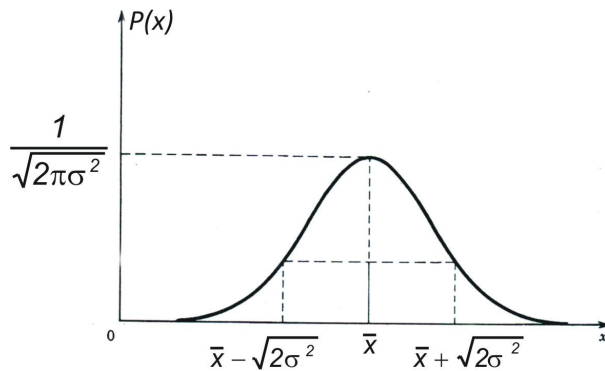


Figura 3. Curva de distribuição de Gauss ou curva normal.

Observa-se que $\frac{1}{\sqrt{2\pi\sigma^2}}$ é o máximo valor da altura (valor mais provável) da função P(x), e que o valor médio \bar{x} representa o valor de “x” para o qual a função P(x) atinge esse valor. Além disso, percebe-se que o termo $\sqrt{2\sigma^2}$ tem a ver com a largura ou estreiteza da curva em forma de boca de sino. Pode-se notar, também, que quanto maior for o expoente da função P(x), mais rapidamente a curva vai caindo para a abscissa, porém, nunca chegando a zero.

Abaixo estão representadas curvas normais com médias M diferentes (a), desvios padrões DP diferentes, porém a mesma média M (b) e com as variações nos dois parâmetros (c).

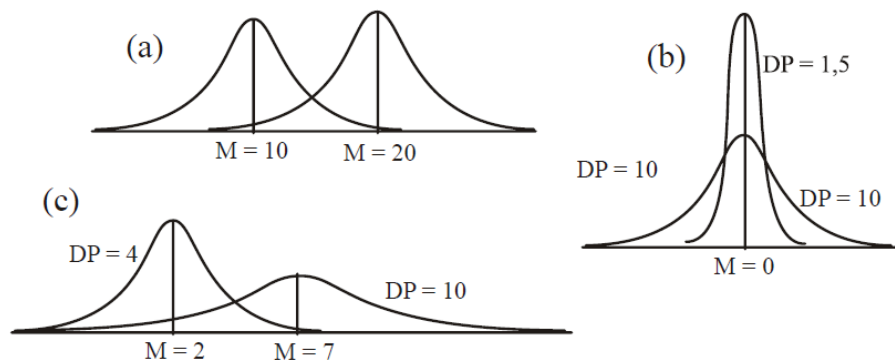


Figura 4. Distribuições normais com diferentes médias e desvios-padrão.

A distribuição de Gauss pode ser usada para se obter a probabilidade de uma medida estar entre quaisquer limites especificados; para isso, basta calcular a área da curva compreendida entre estes limites. Na Figura 5, apresenta-se a representação gráfica desta ideia: o valor da probabilidade de obter para a variável x um valor entre a e b é dada pela área da figura hachurada. Matematicamente, esta área é calculada por

$$\text{meio da integral } P(a,b) = \int_a^b P(x)dx.$$

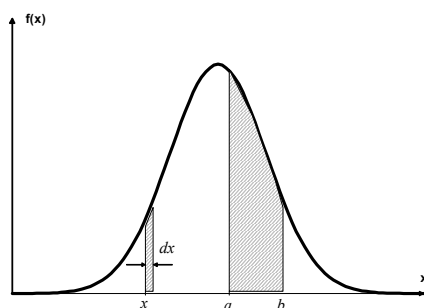


Figura 5. A probabilidade de encontrar o valor da variável entre a e b é dada pela área hachurada na figura do gráfico da distribuição de probabilidade.

Para o caso da distribuição gaussiana, é importante saber o valor da probabilidade de encontrar o resultado entre a média e um ou dois desvio padrão a partir dessa média.

Constata-se, a partir dos cálculos destas áreas, como mostrado na Figura 6, que

- a probabilidade de que a variável assuma um valor entre a média menos um desvio padrão $(\bar{X} - \sigma)$ e a média mais um desvio padrão $(\bar{X} + \sigma)$ é de 68 %;
- para região entre $(\bar{X} - 2\sigma)$ e $(\bar{X} + 2\sigma)$ é de 95,5%,
- e para a região entre $(\bar{X} - 3\sigma)$ e $(\bar{X} + 3\sigma)$ é de 99,7%.

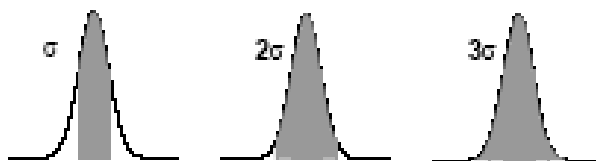


Figura 6. As áreas que representam a probabilidade de encontrar os valores a uma distância de um, dois e três desvios padrão da média.

Em outras palavras, se uma medida tem como resultado uma distribuição no formato da gaussiana, podemos garantir que 99,7% dos resultados obtidos na faixa que dista 3 desvios padrão da média.

Ajuste de dados - Regressão Linear

Quando estamos fazendo medidas em um laboratório, muitas vezes queremos descrever os resultados obtidos em termos de modelos e teorias conhecidas. Às vezes, queremos encontrar uma relação entre as variáveis. Uma das formas de fazer isso envolve o que chamamos de ajuste de dados por uma curva, e a curva mais fácil de se ajustar é uma reta (ou função afim).

Por exemplo, num laboratório fizemos medidas do movimento de um objeto que achamos estar se movendo em movimento uniforme. Obtivemos medidas de posição como função do tempo que estão representadas na Tabela 8.

Tabela 8. Medidas de posição como função do tempo de um objeto em movimento horizontal sem atrito.

t (s)	x (cm)
0,1	1,0
0,2	1,3
0,3	1,5
0,4	1,9
0,5	2,2
0,6	2,4
0,7	2,8
0,8	3,1
0,9	3,3
1,0	3,5
1,1	3,9

Se colocamos esses pontos sobre um gráfico, obtemos o que está apresentado na Figura 7.

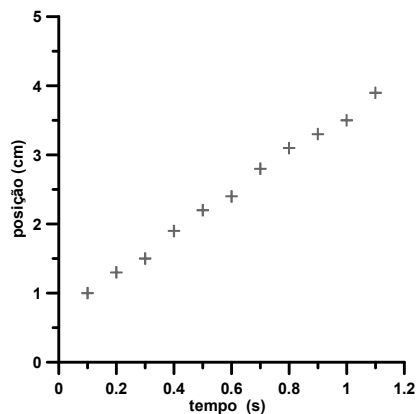


Figura 7. O gráfico da posição como função do tempo para os dados da Tabela 8.

Ao olharmos para este gráfico, observamos nitidamente que os pontos estão alinhados como sobre uma linha reta. Então, é razoável supor que esses dados podem ser bem representados por uma função linear do tipo $f(x)=at+b$. O trabalho, agora, é descobrir qual é a melhor reta que descreve esses dados.

O procedimento para determinação dessa melhor reta é denominado regressão linear. De fato, o que falta é definir um critério para utilizar-se a palavra melhor. Costuma-se usar como critério calcular a diferença quadrática entre o ponto experimental (t_i, x_i) e o valor esperado (t_i, at_i+b), somar as diferenças para todos os pontos e calcular o que chamamos de chi quadrado (χ^2). A melhor reta é aquela que têm os valores de a e b tais que o valor de χ^2 é o mínimo possível. É claro que esse procedimento é matematicamente trabalhoso, mas a maior parte dos softwares gráficos já o faz automaticamente.

Testes: a Teoria Clássica de Testes

Nos processos de avaliação de aprendizagem, precisamos atribuir um escore (uma nota), em geral comparativa, aos estudantes. O que usualmente é feito é atribuir uma pontuação a uma série de questões ou problemas, e somar a pontuação obtida pelo aluno para obter uma nota final. Este mecanismo de obtenção do escore a ser atribuído ao aluno é característica da denominada Teoria Clássica dos Testes (TCT).

Alguns dos problemas enfrentados na Teoria Clássica dos Testes (TCT), quando se pensa em avaliações de larga escala, é a viabilidade de comparação entre os grupos respondentes e a dependência existente entre as características do teste e do examinando. É conhecido de todos os professores que existem provas ou testes “fáceis” ou “difíceis”, turmas “fracas” e “fortes”. Para construir uma forma de elaboração de testes que permita comparações com significados entre grupos e épocas diferentes, utiliza-se a denominada Teoria da Resposta ao Item, que transpõe algumas das limitações observadas na Teoria Clássica de Testes.

Na Teoria Clássica de Testes o foco está na produção de testes de qualidade que no final resultaram em testes válidos. Isso se deve ao fato de que ela se ocupa da explicação do resultado final total do mesmo, ou seja, o escore total que é dado pela soma das respostas corretas de um dado conjunto de itens.

Por exemplo, num questionário de física térmica, tem-se 26 questões a serem respondidas pelos alunos. Um determinado aluno (por exemplo o de número 46), ao respondê-lo, acertou 14 itens. Como foi atribuído grau 0 para cada erro e grau 1 para cada um dos acertos, pode-se afirmar que este aluno, neste teste, apresentou um escore total igual a 14. A TCT, então, se ocupa em dar significado ao que representa esse escore para o referido aluno.

Esta técnica porém apresenta dificuldades bem conhecidas por todos os professores. Para comparar grupos diferentes em momentos diferentes utilizando testes diferentes, a TCT apresenta limitações, das quais algumas são:

1) Os parâmetros clássicos utilizados para analisar um item (uma questão) de um teste (uma prova), a dificuldade do item e a capacidade de discriminação do item, dependem diretamente da amostra de sujeitos utilizada para estabelecê-los. Isto é, é difícil preparar testes ou provas equivalentes para aplicação até mesmo dentro uma escola.

Por exemplo, quando um professor deseja determinar o nível de dificuldade de um teste feito por ele, normalmente ele pede a um colega de sua área para avaliar este teste ou então verifica o resultado do mesmo numa turma que já o tenha feito. No entanto, ao aplicá-lo em outros grupos de alunos, ele percebe que o teste apresenta níveis de dificuldade diferentes, de acordo com as características de cada grupo, ou seja, o resultado da amostra (dos alunos que fizeram o testes) não é válido para análise

das outras turmas que não fizeram o teste. Esta característica da TCT extrapola os muros das instituições de ensino, ocorrendo, por exemplo, nas pesquisas de opinião.

2) O teste que é aplicado interfere diretamente na avaliação das aptidões dos estudantes que realizam este teste. Ao se aplicar testes diferentes para medir a mesma aptidão, obtém-se escores diferentes da mesma aptidão para sujeitos idênticos. Os escores também serão diferentes quando se aplica testes de dificuldades diferentes. No caso das formas paralelas de testes, é preciso observar que, em primeiro lugar, conseguir formas estritamente paralelas é uma tarefa quase impossível e, em segundo lugar, mesmo conseguindo formas paralelas, é difícil pressupor que elas produzam o mesmo montante de erro, o que vem afetar a estimação do escore verdadeiro dos sujeitos.

Observe, por exemplo, os dois itens apresentados na Figura 8, constantes do vestibular da UFRJ, respectivamente nos anos de 1999 e 2004.

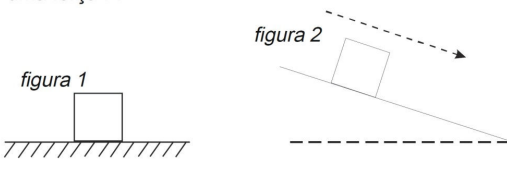
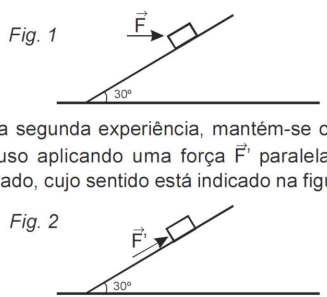
<p>QUESTÃO 2</p> <p>A figura 1 mostra um bloco em repouso sobre uma superfície plana e horizontal. Nesse caso, a superfície exerce sobre o bloco uma força \vec{f}. A figura 2 mostra o mesmo bloco deslizando, com movimento uniforme, descendo uma rampa inclinada em relação à horizontal segundo a reta de maior declive. Nesse caso a rampa exerce sobre o bloco uma força \vec{f}'.</p>  <p>Compare \vec{f} e \vec{f}' e verifique se $\vec{f} < \vec{f}'$, $\vec{f} = \vec{f}'$ ou $\vec{f} > \vec{f}'$. Justifique sua resposta.</p>	<p>QUESTÃO 2</p> <p>Deseja-se manter um bloco em repouso sobre um plano inclinado 30° com a horizontal. Para isso, como os atritos entre o bloco e o plano inclinado são desprezíveis, é necessário aplicar sobre o bloco uma força. Numa primeira experiência, mantém-se o bloco em repouso aplicando uma força horizontal \vec{F}, cujo sentido está indicado na figura 1.</p>  <p>Calcule a razão \vec{F}' / \vec{F}.</p>
--	---

Figura 8. Itens do vestibular da UFRJ 1999 (à esquerda) e 2004 (à direita).

Os dois itens compreendem a aprendizagem de conceitos relacionados à decomposição de grandezas vetoriais e o equilíbrio da partícula. No entanto, no item de 2004, a colocação da força \vec{F} numa direção não comumente utilizada nas resoluções de planos inclinados torna o item mais difícil que o de 1999. Supondo, então, que essas duas questões fossem aplicadas para determinar a aptidão dos alunos em relação a estes conceitos, o fato de a segunda questão ser mais difícil faria com que um mesmo aluno tivesse valores diferentes para sua aptidão dependendo de qual dos dois itens lhe fosse proposto.

Em outras palavras, construir dois testes que apresentem exatamente as mesmas características, avaliem as mesmas aptidões e ainda mais, apresentem a mesma dificuldade para todos os estudantes, afim de que sejam comparáveis, não é uma tarefa possível dentro desta teoria.

Dentro dessas limitações, ao se aplicar um teste, a característica do examinando na qual se está interessado é a “aptidão” medida pelo instrumento de avaliação. A TCT expressa essa “aptidão” pelo escore verdadeiro que é definido como o valor esperado de desempenho observado no teste de interesse. Porém, na determinação desta aptidão na TCT enfrenta-se um grande problema que é o fato de que as características do examinando e do teste não podem ser separadas, ou seja, a aptidão do examinando é definida em termos de um teste particular. Isto é evidenciado quando se aplica um teste considerado difícil ao examinando, que neste caso pode obter um resultado que indique que o ele apresenta uma aptidão baixa. No entanto, se for aplicado outro teste que exija a mesma aptidão, mas este teste possuir um nível de dificuldade menor, pode-se obter, para o mesmo estudante, uma aptidão alta.

Tem-se, então, que na TCT as características do teste e do item mudam quando o contexto do examinando muda, e as características do examinando mudam quando o contexto do teste muda. Portanto, é muito difícil comparar examinandos que fazem testes diferentes, é muito difícil comparar itens cujas características são obtidas usando grupos diferentes de examinandos. Esse problema é constante na atuação dos professores que precisam fazer avaliações diferentes para turmas que estão cursando a mesma série, avaliações substitutas ou de segunda chamada. Por exemplo, um aluno que faz uma prova na data prevista e outro que por algum motivo necessita fazer a segunda chamada da mesma, mesmo tirando a mesma nota, não terão a mesma aptidão.

Testes: a Teoria da Resposta ao Item

Buscando-se superar as limitações existentes na TCT, outras metodologias de avaliação foram propostas e, dentre elas, a Teoria de Resposta ao Item (TRI). Como características inerentes a esta teoria tem-se que:

- 1) os itens (questões) não são dependentes do grupo que faz o teste.
- 2) os escores descrevem a proficiência que independe do teste aplicado.
- 3) esse modelo de avaliação é expresso a partir do item e não do teste como um todo, permitindo que se analise cada item individualmente, independentemente dos demais itens do teste.
- 4) não há exigência de testes estritamente paralelos para a verificação de sua confiabilidade.
- 5) esse modelo fornece uma medida de precisão para cada escore de habilidade.

Nesse modelo de análise, busca-se medir o que se denomina por traço-latente ou variável não-observável. Para uma melhor compreensão do que isso significa, suponha que um professor de física, do ensino médio ou superior, ao avaliar o quanto seus alunos aprenderam (e isso é um conceito abstrato), utiliza como instrumento para obter essa informação uma prova que contém alguns problemas; neste caso, a variável observada por ele é a capacidade de cada estudante resolver os problemas apresentados ou determinado tipo de problema e não há dúvidas que aprender algum tópico de física não é equivalente a saber resolver problemas daquele tópico.

Outro exemplo de medida de traço latente ou variável não observável pode ser encontrado na Matriz de Referência do ENEM¹. Suponha que, ao se construir a prova de Ciências da Natureza, deseja-se medir a competência de se apropriar de conhecimentos da física para, em situações problema, interpretar, avaliar ou planejar intervenções científico-tecnológicas (competência 6), que nesse documento representa uma variável não-observável ou traço latente. Para isso, constróem-se itens que contemplem as habilidades (variável observável) listadas abaixo que, em conjunto, possibilitam determinar essa competência (traço latente).

Tabela 9. Habilidades (variáveis observáveis) pertencentes à competência 6.

Variáveis observáveis ou habilidades referentes à competência 6
H20 - Caracterizar causas ou efeitos dos movimentos de partículas, substâncias, objetos ou corpos celestes.
H21 - Utilizar leis físicas e (ou) químicas para interpretar processos naturais ou tecnológicos inseridos no contexto da termodinâmica e (ou) do eletromagnetismo.
H22 - Compreender fenômenos decorrentes da interação entre a radiação e a matéria em suas manifestações em processos naturais ou tecnológicos, ou em suas implicações biológicas, sociais, econômicas ou ambientais.
H23 - Avaliar possibilidades de geração, uso ou transformação de energia em ambientes específicos, considerando implicações éticas, ambientais, sociais e/ou econômicas.

A Teoria da Resposta ao Item corresponde então a uma proposta psicométrica para determinação de aptidões (traços latentes) que possui dois postulados básicos:

Primeiro Postulado

O desempenho de um examinando num item do teste pode ser previsto (ou explicado) por um conjunto de fatores denominados de traços, traços latentes, ou aptidão.

Segundo Postulado

A relação entre o desempenho em um item e o conjunto de traços que definem este desempenho no item pode ser descrita por uma função monótona-mente crescente da aptidão chamada função característica do item, ou curva característica do item (ICC=item characteristic curve).

Observe que o primeiro postulado traz consigo a idéia básica de que, através da realização de testes, é possível se medir a aptidão do examinando.

O segundo postulado, por sua vez, apresenta uma idéia bastante razoável que consiste no fato de que ao se construir uma curva da probabilidade de acerto de um item para um aluno com determinada aptidão, essa curva deve assumir a forma de um S deitado, isto é, se este aluno possuir uma nota alta, a probabilidade de ele acertar qualquer item é também alta, se este aluno possuir nota média, a probabilidade de acertar qualquer item é média e finalmente, se o aluno tem nota baixa, a probabilidade também será baixa.

¹ Neste caso se utiliza o termo competência como sinônimo de traço latente (aptidão) e o termo habilidade como sinônimo de comportamento ou variável observável.

Considere, por exemplo, que um mesmo teste seja aplicado a duas turmas distintas que possuam níveis de aptidão diferentes, isto é, uma turma com menor nível de aptidão (turma 1) e a outra com maior nível de aptidão (turma 2). Podemos construir uma curva que represente a distribuição dos escores do teste para as duas turmas. Certamente a turma 2 apresentará um valor médio mais alto do que a turma 1. Na Figura 8, apresenta-se um gráfico onde a abscissa corresponde ao escore do aluno, ou melhor, à habilidade ou aptidão do aluno. Na parte inferior do gráfico, estão representadas (de forma invertida) as duas distribuições de notas para cada uma das duas turmas. Na parte superior da curva, representa-se a ICC dos dois grupos, isto é, a curva que representa, para um item do teste, a probabilidade de acerto no item para o candidato cujo escore é dado.

Observe que a ICC para as duas turmas é a mesma, sendo que se pode observar que os respondentes com maior aptidão (localizados mais à esquerda no eixo das abscissas) possuem também maior probabilidade de responder o item corretamente.

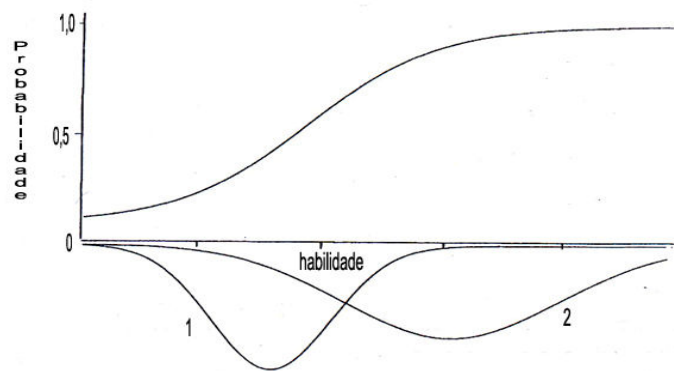


Figura 9. ICC e curvas de distribuição de aptidão para as turmas 1 e 2 de examinandos.

Como na utilização da metodologia da TRI os itens não são dependentes do grupo que é submetido ao teste, os escores que descrevem a aptidão de cada respondente independem do teste aplicado, e os parâmetros estimados em diferentes grupos também devem ser os mesmos, tanto o item quanto os parâmetros de aptidão são ditos invariantes. Isto pode ser observado na figura 3.10, onde dois examinandos que possuem a mesma aptidão, seja na turma 1 ou na turma 2, possuirão a mesma probabilidade de acertar um item.

Os modelos matemáticos que fundamentam a TRI especificam que a probabilidade de um examinando marcar a alternativa correta depende de sua aptidão ou do conjunto formado por esta aptidão e as características do item. A aplicação destes modelos estão alicerçados em três hipóteses:

1. A hipótese de unidimensionalidade do item: uma única aptidão é medida por cada item que compõe o teste. Esta característica é muito difícil de ser verificada, sendo que normalmente basta que exista uma componente ou fator dominante da referida aptidão que influencie no processo.

2. A independência local que ocorre quando as aptidões que influenciam no desempenho no teste são mantidas constantes e as respostas a qualquer par de itens são estatisticamente independentes (ou seja, os únicos fatores que influenciam as respostas aos itens do teste são as aptidões especificadas no modelo)

3. E finalmente, o fato que a função característica do item reflete a relação verdadeira entre as variáveis não observáveis (aptidões) e as variáveis observáveis (respostas aos itens).

Construir uma ICC (curva característica do item) e o modelo de Rasch²

Quando se deseja estimar a aptidão de um atleta de salto em altura, esta pode ser obtida através:

- de um recorde individual;
- de um recorde individual durante um evento oficial ou internacional;
- da média do desempenho do atleta durante um determinado período de tempo; ou
- do desempenho mais freqüente num determinado período de tempo.

Suponha, então, que para realização dessa tarefa, tomou-se os resultados obtidos por dois atletas de salto em altura durante um ano e construiu-se um diagrama da probabilidade de sucesso desses atletas em função da altura saltada.

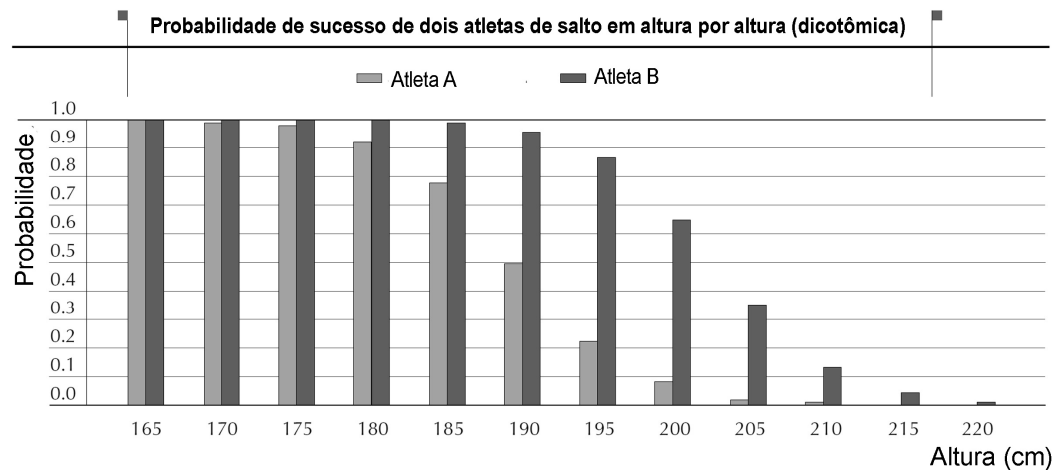


Figura 10. Probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano.

Observe que os dois atletas sempre têm sucesso em 165 cm. A partir daí, a probabilidade de sucesso cai até alcançar 0 (zero) para ambos quando a altura a ser saltada é de 225 cm. Para o primeiro atleta, essa diminuição da probabilidade se inicia a altura de 170 cm, enquanto para o segundo, o mesmo fato só ocorre a partir de 185 cm.

Estas informações podem ser tratadas com o modelo de regressão logística. Esta análise estatística consiste em transformar a variável dicotômica (sim/não, sucesso/fracasso, 1/0) em uma variável contínua. Neste exemplo, a variável contínua indicará o sucesso ou fracasso de um atleta em particular em função da altura do salto. Os resultados desta análise permitirão estimar a probabilidade de sucesso em qualquer

² Exemplo retirado de PISA 2009 Data Analysis Manual, OECD.
http://www.oecd.org/document/19/0,3746,en_2649_35845621_48577747_1_1_1_1,00.html

altura. O diagrama da Figura 11 apresenta a possibilidade de sucesso para os dois atletas.

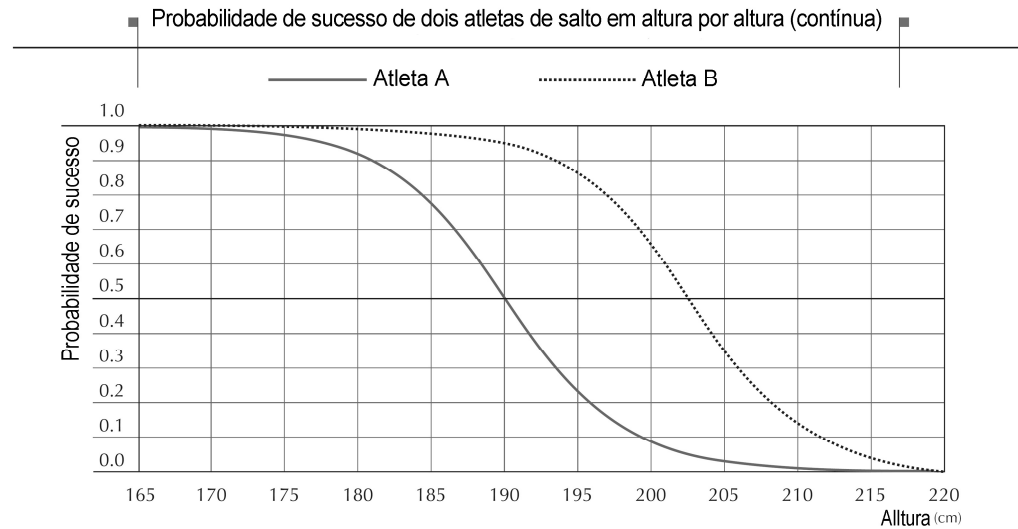


Figura 11. Probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano.

Estas duas curvas representam a probabilidade de sucesso para os dois atletas. A curva sólida representa a probabilidade de sucesso do atleta A e pontilhada do atleta B. Por definição, o nível de desempenho pode ser definido como a altura em que a probabilidade de sucesso é 0,5. Isso faz sentido uma vez que abaixo deste nível a probabilidade de sucesso passa a ser menor do que a de fracasso e acima dele, o inverso.

Neste exemplo em particular, o nível de desempenho dos dois atletas é respectivamente 190 cm e 202,5 cm. Observe na Figura 11 que o nível de desempenho do atleta A é visto diretamente no gráfico, enquanto o do atleta B é estimado do modelo. Uma propriedade fundamental deste tipo de abordagem é que o nível (ou seja, a altura) do sarrafo a ser saltado e o desempenho dos atletas são expressas na mesma medida ou escala.

A mesma idéia até aqui explorada está por trás do modelo de Rasch (um tipo de modelo logístico de um parâmetro) para a TRI. A dificuldade dos itens em um teste é análoga à dificuldade do salto com base na altura da barra. Além disso, assim como um salto em particular possui dois resultados possíveis (sucesso ou fracasso), a resposta de um aluno a um determinado item também possui duas possibilidades (acerto ou erro). Finalmente, assim como o desempenho do atleta foi definida no ponto onde a probabilidade de sucesso é de 0,5, o desempenho ou aptidão do aluno é medida no ponto onde sua probabilidade de sucesso no item é de 0,5.

Uma das características importantes do modelo de Rasch é que ele cria um contínuo no qual tanto o desempenho do estudante quanto a dificuldade do item estarão localizados em uma função probabilística que relaciona essas duas componentes. Alunos de baixo desempenho ou aptidão e itens fáceis estão localizados na parte inferior do contínuo ou escala enquanto alunos de alto desempenho ou aptidão e itens difíceis estarão localizados na parte superior do contínuo ou escala. Na Figura 12, está representada a probabilidade de sucesso e de fracasso num item de dificuldade zero.

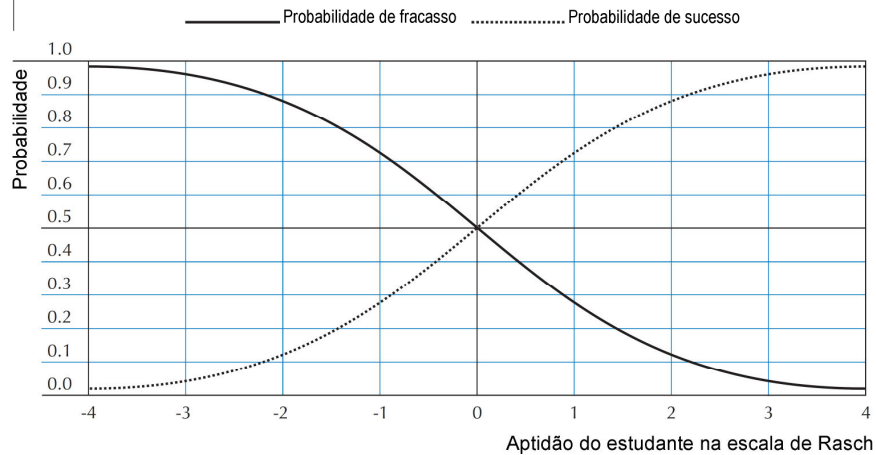


Figura 12. Probabilidade de sucesso de dois atletas de salto em altura pela altura saltada no período de um ano.

Como se pode observar acima, um estudante com uma aptidão zero tem a probabilidade de 0,5 de sucesso ou fracasso num item de dificuldade zero. Um estudante com aptidão - 2 tem a probabilidade de um pouco mais de 0,10 de sucesso e um pouco menos de 0.90 de fracasso no mesmo item de dificuldade zero. Mas, este último, teria probabilidade 0,5 de sucesso num item de dificuldade - 2.

Do ponto de vista matemático, a probabilidade de um estudante com uma aptidão “ θ ” responder corretamente um item de dificuldade “ b ” é:

$$\text{"odds ratio"} = \frac{\text{probabilidade (acerto)}}{\text{probabilidade (fracasso)}} = \frac{P}{1-P}$$

$$\frac{P}{1-P} = e^{(\theta-b)} \quad \rightarrow \quad P(\theta) = \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} = \frac{1}{1+e^{-(\theta-b)}}$$

De forma similar, a probabilidade de fracasso $P(\theta) = 0$ é dada por:

$$P(\theta_{fracasso}) = 1 - \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} \quad \rightarrow \quad P(\theta_{fracasso}) = \frac{1+e^{(\theta-b)} - e^{(\theta-b)}}{1+e^{(\theta-b)}} = \frac{1}{1+e^{(\theta-b)}}$$

Observe que $P(\theta) + P(\theta_{fracasso}) = 1$, ou seja:

$$P(\theta) + P(\theta_{fracasso}) = \frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} + \frac{1}{1+e^{(\theta-b)}} = \frac{1+e^{(\theta-b)}}{1+e^{(\theta-b)}} = 1$$

Em outras palavras, considera-se que a curva característica do item pode ser ajustada (no sentido discutido antes, de regressão) por uma curva do tipo mostrado para $P(\theta)$.

A seguir, apresentam-se alguns exemplos do cálculo da probabilidade de sucesso para aptidões e níveis de dificuldades estabelecidos.

1º exemplo) Cálculo da probabilidade de sucesso quando a habilidade do estudante é igual a dificuldade do item.

Tabela 10. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é igual a dificuldade do item.

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-2	0,5
-1	-1	0,5
0	0	0,5
1	1	0,5
2	2	0,5

$$\theta - b = 0 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(0)}} = \frac{1}{1+1} = 0,5$$

Logo, a probabilidade de fracasso será: $P(\theta \text{ fracasso}) = 1 - 0,5 = 0,5$

2º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é uma unidade menor do que a dificuldade do item.

Tabela 11. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é uma unidade menor do que a dificuldade do item.

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-1	0,27
-1	0	0,27
0	1	0,27
1	2	0,27
2	3	0,27

$$\theta - b = -1 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(-1)}} = \frac{1}{1 + (2,718)^1} = 0,27$$

Logo, a probabilidade de fracasso será: $P(\theta \text{ fracasso}) = 1 - 0,27 = 0,73$

3º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é uma unidade maior do que a dificuldade do item.

Tabela 12. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é uma unidade maior do que a dificuldade do item.

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-3	0,73
-1	-2	0,73
0	-1	0,73
1	0	0,73
2	1	0,73

$$\theta - b = 1 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(1)}} = \frac{1}{1 + (2,718)^{-1}} = 0,73$$

Logo, a probabilidade de fracasso será: $P(\theta \text{ fracasso}) = 1 - 0,73 = 0,27$

4º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é duas unidades menor do que a dificuldade do item.

Tabela 13. Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é duas unidades menores do que a dificuldade do item.

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	0	0,12
-1	1	0,12
0	2	0,12
1	3	0,12
2	4	0,12

$$\theta - b = -2 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-(-2)}} = \frac{1}{1 + (2,718)^2} = 0,12$$

Logo, a probabilidade de fracasso será: $P(\theta \text{ fracasso}) = 1 - 0,12 = 0,88$

5º exemplo) Cálculo da probabilidade de sucesso quando a aptidão do estudante é duas unidades maior do que a dificuldade do item.

Tabela 14 - Relação de aptidão do estudante, dificuldade do item e probabilidade de sucesso quando a aptidão do estudante é duas unidades maior do que a dificuldade do item.

Aptidão do estudante (θ)	Dificuldade do item (b)	Probabilidade de sucesso
-2	-4	0,12
-1	-3	0,12
0	-2	0,12
1	-1	0,12
2	0	0,12

$$\theta - b = 2 \rightarrow P(\theta) = \frac{1}{1 + e^{-(\theta-b)}} = \frac{1}{1 + e^{-2}} = \frac{1}{1 + (2,718)^{-2}} = 0,88$$

Logo, a probabilidade de fracasso será: $P(\theta \text{ fracasso}) = 1 - 0,88 = 0,12$

Observe que:

* se a aptidão do estudante é igual a o nível de dificuldade do item, a probabilidade de sucesso será sempre igual a 0,5, independentemente da posição da aptidão e da dificuldade do item no contínuo;

* se a dificuldade do item excede aptidão do estudante em uma unidade, a probabilidade de sucesso será sempre igual a 0,27, independentemente da posição da aptidão e da dificuldade do item no contínuo;

* se aptidão do estudante excede a dificuldade do item em uma unidade, a probabilidade de sucesso será sempre igual a 0,73, independentemente da posição da aptidão e da dificuldade do item no contínuo;

* se duas unidades separam a aptidão do estudante e o nível de dificuldade do item, a probabilidade de sucesso será igual a 0,12 e 0,88, respectivamente, independentemente da posição da aptidão e da dificuldade do item no contínuo.

Desses exemplos, fica evidente que somente a distância entre a aptidão do estudante e o nível de dificuldade do item, no contínuo de Rasch, que será fator determinante da probabilidade de sucesso deste aluno na execução de um determinado item.

Os exemplos também ilustram a simetria das escalas. Se um estudante possui aptidão uma unidade abaixo do nível de dificuldade do item, a probabilidade de sucesso deste será de 0,27, ou seja, 0,23 abaixo da probabilidade de sucesso quando essa aptidão e dificuldade do item são iguais. Agora, se ao contrário, o estudante possuir aptidão uma unidade acima do nível de dificuldade do item, sua probabilidade de acerto será de 0,73, ou seja, 0,23 acima da probabilidade de sucesso quando essa aptidão é igual ao nível de dificuldade do item. Da mesma forma, a diferença é de duas unidades gerará um fator de simetria igual a 0,38.

Modelos para a Teoria da Resposta ao Item

Quando falamos de modelo, estamos pensando em ajustar os dados obtidos por uma curva. No caso da Teoria da Resposta ao Item, esse ajuste deve ser feito de forma a fornecer uma função que descreva adequadamente o comportamento da curva característica do item, na forma de um S deitado de lado.

Os modelos existentes para a TRI diferenciam-se basicamente pelo número de parâmetros utilizados para descrever os itens. Os modelos mais populares são os modelos logísticos de um, dois e três parâmetros, sendo que estes são apropriados para dados de respostas de itens dicotômicos.

Modelo Logístico de um Parâmetro

Este modelo para a TRI é um dos mais utilizados. Sua curva característica é dada por

$$P_i(\theta) = \frac{1}{1 + e^{-y(\theta)}} = \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = \theta - b_i \text{ com } i = 1, 2, \dots, n,$$

Nesta expressão,

i refere-se a um item de um total de n itens que constituem o teste; esses itens são dicotômicos, isto é, só admitem respostas do tipo certo (1) ou errado (0); $P_i(\theta)$ é a probabilidade que um examinando, escolhido aleatoriamente, com aptidão θ (parâmetro do modelo), responda ao item i corretamente; a probabilidade tem sempre um valor entre 0 e 1;

b_i é um parâmetro do modelo, que indica o nível de dificuldade do item.

Observa-se que $y(\theta) = \theta - b_i$ corresponde a uma função linear, com coeficiente angular igual a 1. Pode-se verificar que, quando $y(\theta = b_i) = 0$, a probabilidade de acerto $P_i(\theta = b_i) = 0,5$, ou seja, b_i define a aptidão para a qual o candidato tem probabilidade 50% de acertar o item.

Essa parametrização corresponde ao modelo de Rasch, sendo que esse modelo logístico para a TRI é o modelo utilizado pelo PISA.

Veja agora, um exemplo da curva característica do item, da questão intitulada “Clareza” (que se constitui de dois itens), integrante da prova de Ciências, na edição de 2000.

A questão “Clareza” foi tornada pública pelo PISA3 e seu texto é apresentado a seguir. Como em todas as unidades do PISA, uma “questão” é constituída de um texto

³ As questões públicas do PISA podem ser consultadas na página do INEP: <http://portal.inep.gov.br/internacional-novo-pisa-itens>. consultado em

introdutório seguido de um ou mais itens. A unidade Claridade tem dois itens associados. Na figura 3.14 está apresentado o texto introdutório.

CLARIDADE

Leia as informações abaixo e responda às questões que se seguem.

DURAÇÃO DO DIA EM 22 DE JUNHO DE 1998

Hoje, enquanto o Hemisfério Norte celebra seu dia mais longo, os australianos viverão o seu dia mais curto.

Em Melbourne*, Austrália, o sol nascerá às 7:36 h e se porá às 17:08 h, totalizando nove horas e 32 minutos de claridade.

Compare o dia de hoje com o dia mais longo do ano no Hemisfério Sul, esperado para 22 de dezembro, quando o sol nascerá às 5:55 h e se porá às 20:42 h (horário de verão), totalizando 14 horas e 47 minutos de claridade.

O presidente da Sociedade de Astronomia, Sr. Perry Vlahos, disse que a existência de mudanças nas estações entre os Hemisférios Norte e Sul estava ligada à inclinação de 23 graus da Terra.

*Melbourne é uma cidade no sul da Austrália a uma latitude de cerca de 38 graus ao Sul do equador.

Figura 12. Texto introdutório da questão “Claridade” do PISA 2000 [PISA 2000].

O primeiro item apresentado é uma questão de múltipla escolha referente ao conhecimento de fenômenos astronômicos básicos, no caso, especificamente a explicação para a existência de dia e noite (claridade e escuridão). Seu texto está apresentado na Figura 13.

QUESTÃO 24: CLARIDADE S129Q01

Qual é a afirmação que explica por que a claridade e a escuridão ocorrem na Terra?

- A A Terra gira em torno do seu eixo.
- B O Sol gira em torno do seu eixo.
- C O eixo da Terra é inclinado.
- D A Terra gira em torno do Sol.

Figura 13. Texto do item 1 da questão “Claridade” do PISA 2000 [PISA 2000].

A construção das ICC's deste item pode ser feita através da utilização dos dados brutos do PISA (disponíveis em www.pisa.oecd.org). Considerando-se o escore bruto dos participantes do exame como sendo a aptidão, pode-se reconstruir as ICC's para diversas sub-amostras dos participantes. Na Figura 14 apresentamos as ICC's para a item “Claridade - Q1” da figura 3.15 para todos os participantes do PISA, do Brasil, do Japão e do Reino Unido.

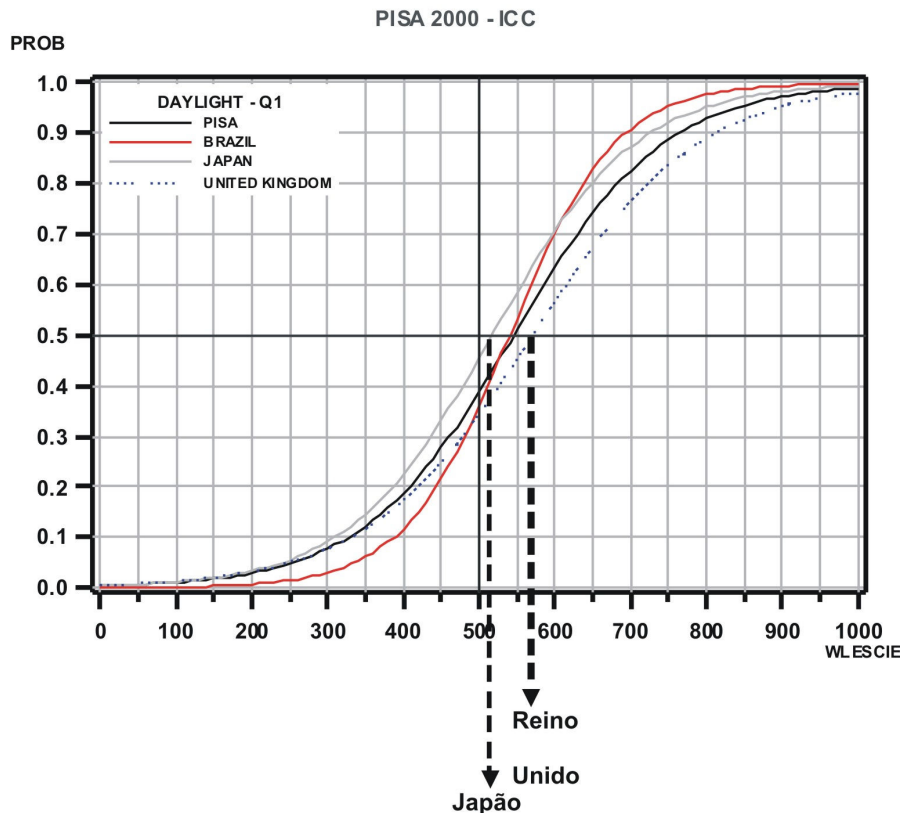


Figura 14. ICC de Todos participantes do PISA 2000, do Brasil, do Japão e do Reino Unido.

Observa-se, pela posição em relação ao eixo de aptidão (WLESCIE no gráfico) que esse item foi um pouco mais difícil para os estudantes do Reino Unido (linha tracejada azul) do que para os do Japão (linha contínua cinza), já que a aptidão que corresponde a uma probabilidade de acerto do item igual a 50% foi maior para o Reino Unido do que para o Japão. O índice de dificuldade para os participantes brasileiros é intermediário entre esses dois países e praticamente igual ao da amostra global.

Esta constatação revela a idéia básica da construção das ICC's. O desempenho dos participantes brasileiros (no PISA 2000) revela-se bastante pior do que do restante do mundo quando observamos os histogramas dos escores no exame. O índice de acerto dos estudantes brasileiros neste item é muito mais baixo do que do total de participantes e dos participantes do Japão e Reino Unido, por exemplo⁴. Mesmo assim, as ICC's são praticamente idênticas.

Na Figura 15 mostramos o item 2 da unidade "Clareza". Este item exige o desenho das posições do eixo da Terra e do Equador, sendo considerada bastante difícil.

⁴ Esses gráficos e resultados estão apresentados em artigos de M.F. Barroso.

QUESTÃO 25: CLARIDADE S129Q02- 0 1 2 8 9

A Figura 1 demonstra os raios de luz do sol se refletindo sobre a Terra.

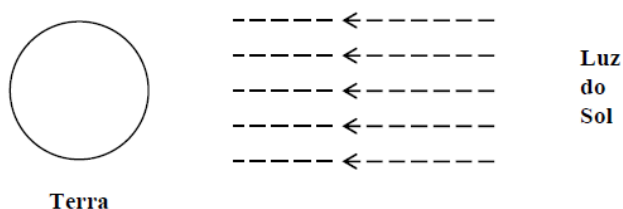


Figura 1

Suponha que seja o dia mais curto em Melbourne. Mostre o eixo da Terra, o Hemisfério Norte e o Hemisfério Sul na Figura 1.

Figura 15. Texto do item 1 da questão “Claridade” do PISA 2000.

A construção das ICC’s deste item para todos os participantes do PISA e para os participantes do Brasil, Japão, Portugal e Reino Unido revela que este item apresenta um grande nível de dificuldade. Como se pode observar na Figura 16, o índice de dificuldade desse item é de 650, ou seja, apenas os participantes cujo escore foi acima da média em 1,5 desvio padrão (menos de $\cong 10\%$ dos participantes) tem probabilidade de acertar este item superior a 50%.

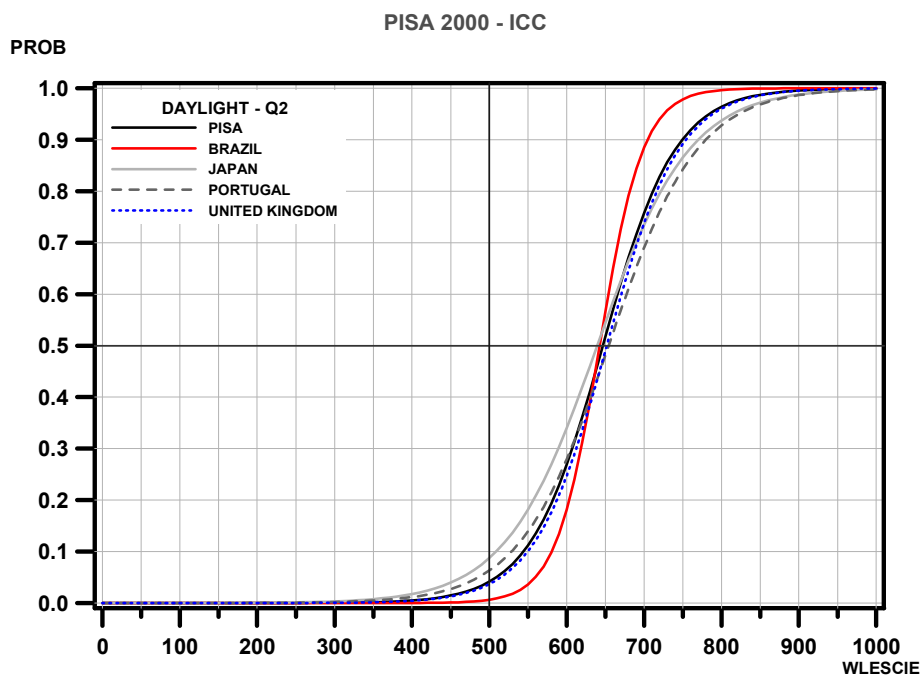


Figura 16. ICC de Todos participantes do PISA 2000, do Brasil, do Japão, de Portugal e do Reino Unido.

Ainda se pode observar na figura o índice de dificuldade deste item 2 é praticamente o mesmo para todos os países representados no diagrama.

Modelo logístico de dois parâmetros

O modelo logístico de dois parâmetros permite avaliar dois parâmetros do item: a dificuldade e a discriminação. A curva característica deste modelo para TRI é dada pela equação

$$P_i(\theta) = \frac{1}{1 + e^{-y(\theta)}} = \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = Da_i(\theta - b_i) \text{ com } i = 1, 2, \dots, n,$$

onde

D é um fator de escala introduzido em virtude de considerações estatísticas com valor 1,75;

a_i representa o parâmetro de discriminação do item, proporcional à inclinação da

ICC no ponto b_i na escala de aptidão e é determinado por $\left. \frac{dP_i}{d\theta} \right|_{\theta=b_i} = \frac{D}{4} a_i$

Esse valor pode variar de 0 a ∞ , mas, normalmente, varia entre 0 e 2. Valores negativos não são aceitos pois representariam que quanto menor a aptidão do estudante, maior a possibilidade dele acertar o item.

A Figura 17 mostra os parâmetros de dificuldade (b_i) e de discriminação (a_i) de dois itens.

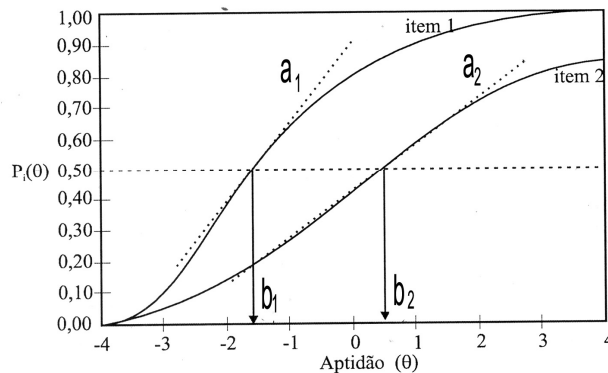


Figura 17. Parâmetros de dificuldade (b) e discriminação (a) de dois itens.

Observe que o item 2 apresenta uma maior nível de dificuldade que o item 1 (valor de b maior). No entanto, o item 1 é mais discriminatório, pois possui maior inclinação no ponto correspondente a probabilidade de acerto igual a 0,5.

Modelo logístico de três parâmetros

O modelo logístico de três parâmetros permite avaliar dois parâmetros do item: a dificuldade, a discriminação e a resposta dada ao acaso (o chute). A curva característica deste modelo para TRI é dada pela equação

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{y(\theta)}}{1 + e^{y(\theta)}}, \text{ sendo } y(\theta) = Da_i(\theta - b_i) \text{ com } i = 1, 2, \dots, n,$$

onde “ c_i ” é o parâmetro do item que permite avaliar a resposta correta dada ao item por acaso e é expresso pela assíntota inferior da curva, no nível mais baixo do contínuo de aptidões.

Na Figura 18, apresentam-se os parâmetros de dificuldade (b_i), de discriminação (a_i) e de “chute” (c_i) de três itens.

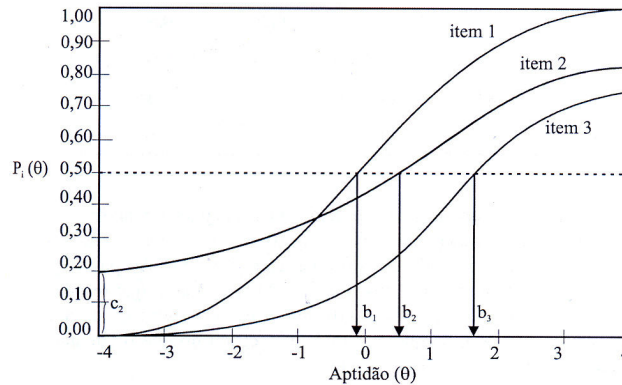


Figura 18. Parâmetros de dificuldade (b) e discriminação (a) de dois itens.

Observe que existe um nível de dificuldade crescente, do item 1 para o 3, que pode ser observado pela posição das ICC (quando $P_i(\theta) = 0,5$) em relação ao eixo das aptidões ($b_1 < b_2 < b_3$). É também possível determinar o índice de discriminação de cada uma das curvas, pela inclinação da mesma nesse mesmo ponto (quando $P_i(\theta) = 0,5$).

O que difere o modelo de três parâmetros é que, no item 2 por exemplo, o ponto da ordenada que é cortado pela ICC indica a existência de chute, cuja probabilidade nesse item é de 20%. Observe que nos itens 1 e 2 essa probabilidade é zero. Segundo Pasquali,

“A lógica que fundamenta essa interpretação da assíntota é a seguinte: supostamente o sujeito não tem habilidade praticamente nenhuma, pois ele tem um θ menor que -3, e apesar disso acerta o item; conseqüentemente, ele só pode ter chutado e teve sorte, porque acertou.”

O modelo de três parâmetros é o adotado para análise das provas do Exame Nacional do Ensino Médio

Após essa discussão, pode-se observar que a Teoria de Resposta ao Item e a ideia de construção de curvas características do item constituem-se em poderosas ferramentas para a análise de avaliações de aprendizagem. A TRI, devido à independência do teste em relação ao examinando, permite a avaliação e a comparação de diferentes grupos com diferentes aptidões. Isso é de fundamental importância para que se possa obter resultados confiáveis que permitam a tomada de ações sejam pelos professores ou elaboradores de políticas públicas.