



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Tecnologia
Câmpus de Presidente Prudente

Estudo sobre Construção de Escalas com Base na Teoria da Resposta ao Item: Avaliação de Proficiência em Conteúdos Matemáticos Básicos.

Tânia Robaskiewicz Coneglian Fujii

Orientadora: Prof. Dra. Aparecida Donizete Pires de Souza

Coorientadora: Prof. Dra. Mariana Curi

Programa: Matemática Aplicada e Computacional

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências e Tecnologia de Presidente Prudente

Programa de Pós-Graduação em Matemática Aplicada e Computacional

**Estudo sobre Construção de Escalas com Base
na Teoria da Resposta ao Item: Avaliação de
Proficiência em Conteúdos Matemáticos
Básicos.**

Tânia Robaskiewicz Coneglian Fujii

Orientadora: Prof. Dra. Aparecida Donizete Pires de Souza

Coorientadora: Prof. Dra. Mariana Curi

Dissertação apresentada ao Programa de Pós-graduação em Matemática Aplicada e Computacional da Faculdade de Ciências e Tecnologia da UNESP, para a obtenção do título de mestre em Matemática Aplicada e Computacional, sob orientação da Prof. Dra. Aparecida Donizete Pires de Souza.

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação - Diretoria Técnica de Biblioteca e Documentação - UNESP, Campus de Presidente Prudente

F971e Fujii, Tânia Robaskiewicz Coneglian.
Estudo sobre construção de escalas com base na Teoria da Resposta ao Item : avaliação de proficiência em conteúdos matemáticos básicos / Tânia Robaskiewicz Coneglian Fujii. - 2018
141 f.

Orientador: Aparecida Donizete Pires de Souza
Coorientador: Mariana Curi
Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Ciências e Tecnologia, Presidente Prudente, 2018
Inclui bibliografia

1. Teoria da Resposta ao Item. 2. Proficiência em conteúdos matemáticos básicos. 3. Inferência Bayesiana. 4. Segmentação da escala. I. Souza, Aparecida Donizete Pires de. II. Curi, Mariana. III. Universidade Estadual Paulista. Faculdade de Ciências e Tecnologia. IV. Título.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO: Estudo sobre Construção de Escalas com Base na Teoria da Resposta ao Item:
Avaliação de Proficiência em Conteúdos Matemáticos Básicos

AUTORA: TÂNIA ROBASKIEWICZ CONEGLIAN FUJII

ORIENTADORA: APARECIDA DONIZETE PIRES DE SOUZA

COORIENTADORA: MARIANA CURI

Aprovada como parte das exigências para obtenção do Título de Mestra em MATEMÁTICA
APLICADA E COMPUTACIONAL, pela Comissão Examinadora:

Profa. Dra. APARECIDA DONIZETE PIRES DE SOUZA
Departamento de Estatística / Faculdade de Ciências e Tecnologia de Presidente Prudente

Prof. Dr. ADRIANO FERRETI BORGATTO
Departamento de Informática e Estatística / Universidade Federal de Santa Catarina

Prof. Dr. MARIO HISSAMITSU TARUMOTO
Departamento de Estatística / Faculdade de Ciências e Tecnologia de Presidente Prudente

Convidada:

VIDEOCONFERÊNCIA

Profa. Dra. MARIANA CURI - Coorientadora
Instituto de Ciências Matemáticas e de Computação / Universidade de São Paulo

Presidente Prudente, 07 de maio de 2018.

Ao meu querido e amado tesouro, Heliomar.

Agradecimentos

Agradeço em primeiro lugar a Deus, por me fazer forte e perseverante e me manter sempre firme nas horas difíceis.

A todos os meus familiares, meus pais, meu irmão, meus cunhados, cunhadas, sobrinhos, minha sogra, pelo apoio incondicional e compreensão pelas minhas ausências por conta da dedicação aos estudos.

Agradeço especialmente ao meu querido esposo, meu “tesouro”, por sonhar meus sonhos, pelo apoio, incentivo e compreensão sempre.

A todos os professores que tive durante o mestrado, pelo aprendizado, pelo apoio e incentivo. Agradeço em especial a minha orientadora professora Dra Aparecida Donizete Pires de Souza, pela orientação, compreensão e por acreditar em mim.

A professora Monica Furkotter, por toda sua contribuição para o trabalho. A todos os professores do departamento de Matemática e Computação da FCT/UNESP que também contribuíram.

A professora Dra Mariana Curi, pela colaboração como coorientadora neste trabalho.

Ao professor Dr Adriano Ferreti Borgatto, por sua grande colaboração, pelas sugestões e por ser sempre tão gentil e disposto a ajudar.

Agradeço a todos os amigos que fiz durante essa jornada, em especial as “meninas do posMAC”, minhas queridas Clícia, Débora, Karlla e Thais, amigas que vou levar para sempre em meu coração, mesmo que nossos caminhos nos distanciem umas das outras.

Ao meu amigo Enrico, que sempre me deu incentivo, acreditando que sou capaz mesmo eu muitas vezes não acreditando.

A CAPES pelo apoio financeiro.

“Lute com determinação, abrace a vida com paixão, perca com classe e vença com ousadia, porque o mundo pertence a quem se atreve e a vida é muito bela para ser insignificante.”

Charles Chaplin

Resumo

Neste trabalho realizou-se um estudo sobre construção de escalas, com base na Teoria da Resposta ao Item (TRI), resultando na construção e interpretação pedagógica de uma escala de conhecimento para medir a proficiência em conteúdos matemáticos, necessários para o acompanhamento das disciplinas de cálculo e similares dos ingressantes nos cursos da área de exatas. O modelo matemático adotado nesta pesquisa foi o logístico unidimensional de três parâmetros. A estimação dos parâmetros dos itens e das proficiências dos respondentes foi feita sob enfoque bayesiano, utilizando-se o amostrador de Gibbs, algoritmo da classe dos Métodos de Monte Carlo via Cadeia de Markov (MCMC), implementado via software OpenBUGS (Bayesian inference Using Gibbs Sampling), direcionado para análise bayesiana de modelos complexos. O software BILOG-MG também foi utilizado para comparação dos resultados. O instrumento utilizado para a medida do conhecimento consistiu em uma prova composta por trinta e seis itens de múltipla escolha, cada um com cinco alternativas, sendo somente uma a correta. Os itens foram elaborados com base em uma matriz de referência construída para este fim, dividida em três temas, sendo estes “espaço e forma”, “grandezas e medidas” e “números e operações/álgebra e funções”. Cada tema é composto por competências e cada competência descreve uma habilidade que se deseja medir. Para a construção da escala proposta, optou-se por adotar uma escala com média 250 e desvio padrão 50. Nesta escala foram selecionados níveis para serem interpretados em um intervalo de 75 a 425. Para interpretação da escala proposta, foram comparados alguns métodos de posicionamento de itens âncora nos níveis selecionados. Buscando a interpretação da escala, em toda a sua amplitude, optou-se por utilizar a análise de agrupamentos hierárquicos para segmentar a escala em grupos, ou seja, em faixas de proficiência. A escala foi dividida em cinco grupos, cada grupo caracterizado com base nos itens posicionados como âncora, a partir de suas probabilidades de resposta correta e de seus valores para o parâmetro de discriminação. Embora os resultados sejam consistentes, apontam para a necessidade de um processo contínuo de aprimoramento do banco de questões e da escala de proficiência.

Palavras-Chave: *Teoria da Resposta ao Item; Proficiência em Conteúdos Matemáticos Básicos; Inferência Bayesiana; Segmentação da Escala.*

Abstract

In this work, a study was carried out on the construction of scales, based on the Item Response Theory (IRT), resulting in the construction and pedagogical interpretation of a scale of knowledge to measure the proficiency in mathematical contents, necessary for the follow-up of Calculus and similar subjects of the students in the courses of the Exact Sciences Area. The mathematical model adopted in this research was the three parameters one-dimensional logistic. The parameters estimation of the items and proficiencies of the respondents was done using a Bayesian approach using the Gibbs sampler, Monte Carlo Methods via Markov Chain algorithm (MCMC), implemented using OpenBUGS software (Bayesian inference Using Gibbs Sampling), directed to Bayesian analysis of complex models. The BILOG-MG software was also used to compare the results. The instrument used for the measurement of knowledge consisted of a test composed of thirty-six multiple choice items, each with five alternatives, with only one correct. The items were elaborated based on a reference matrix constructed for this purpose, divided in three themes, being these “space and form”, “quantities and measures” and “numbers and operations/ algebra and functions”. Each subject is composed of competencies and each competency describes a skill that one wishes to measure. In order to construct the proposed scale, we chose to adopt a scale with a mean of 250 and standard deviation of 50. In this scale, we selected levels to be interpreted in a range of 75 to 425. For the interpretation of the proposed scale, some methods of positioning anchor items at the selected levels were compared. In order to interpret the scale in all its amplitude, it was decided to use hierarchical groupings analysis to segment the scale into groups, that is, in skill bands. The scale was divided into five groups, each group was characterized based on the items positioned as anchor, from their correct response probabilities and their values for the discrimination parameter. Although the results are consistent, they point to the need for an ongoing upgrading process of questions bank and proficiency scale.

Keywords: Item Response Theory; Proficiency in Basic Mathematical Content; Bayesian Inference; Scale Segmentation. .

Lista de Figuras

3.1	Exemplo de uma curva característica do item - CCI.	41
3.2	Exemplos de curva característica do item com valores diferentes para o parâmetro de discriminação.	42
3.3	Exemplos de curva característica do item com valores diferentes para o parâmetro de dificuldade.	42
4.1	Relação entre o número de acertos e número de indivíduos.	76
4.2	CCI's para os 34 itens considerados.	78
4.3	Curvas características dos itens 1 e 25.	79
4.4	Curvas características dos itens 3 e 36.	79
4.5	Curvas de informação dos itens.	80
4.6	Curva de informação e curva característica do item 36.	80
4.7	Curva de informação do teste.	81
4.8	Relação entre números de acertos e proficiências estimadas.	81
4.9	Posicionamento dos itens âncora nos níveis âncora da escala representados por régua, de acordo com o sugerido por Beaton e Allen (1992) e baseados nos procedimentos utilizado pelo SAEB e SARESP.	86
4.10	Dendograma.	88
D.1	Curva característica e curva de informação do item 2	119
D.2	Curva característica e curva de informação do item 3	120
D.3	Curva característica e curva de informação do item 4	120
D.4	Curva característica e curva de informação do item 5	121
D.5	Curva característica e curva de informação do item 6	121
D.6	Curva característica e curva de informação do item 7	122
D.7	Curva característica e curva de informação do item 8	122
D.8	Curva Característica e curva de informação do item 9	123
D.9	Curva característica e curva de informação do item 10	123
D.10	Curva característica e curva de informação do item 11	124
D.11	Curva característica e curva de informação do item 12	124
D.12	Curva característica e curva de informação do item 13	125
D.13	Curva característica e curva de informação do item 14	125
D.14	Curva característica e curva de informação do item 15	126
D.15	Curva característica e curva de informação do item 16	126
D.16	Curva característica e curva de informação do item 17	127
D.17	Curva característica e curva de informação do item 18	127
D.18	Curva característica e curva de informação do item 19	128
D.19	Curva característica e curva de informação do item 20	128
D.20	Curva característica e curva de informação do item 21	129
D.21	Curva característica e curva de informação do item 22	129

D.22 Curva característica e curva de informação do item 23	130
D.23 Curva característica e curva de informação do item 24	130
D.24 Curva característica e curva de informação do item 25	131
D.25 Curva característica e curva de informação do item 26	131
D.26 Curva característica e curva de informação do item 27	132
D.27 Curva característica e curva de informação do item 28	132
D.28 Curva característica e curva de informação do item 29	133
D.29 Curva característica e curva de informação do item 31	133
D.30 Curva característica e curva de informação do item 32	134
D.31 Curva característica e curva de informação do item 33	134
D.32 Curva característica e curva de informação do item 34	135
D.33 Curva característica e curva de informação do item 35	135
D.34 Curva característica e curva de informação do item 36	136

Lista de Quadros

4.1	Relação dos itens que foram acrescentados e modificados para atual proposta de prova.	49
4.2	Descrição da escala.	88
2.1	Tema I: Espaço e Forma.	133
2.2	Tema II: Grandezas e Medidas.	133
2.3	Tema III: Números e Operações/ Álgebra e Funções.	133

Lista de Tabelas

1.1	Nível de Proficiência em Matemática dos alunos do 3º ano do Ensino Médio (SARESP 2016).	23
4.1	Número mínimo, média e máximo de acertos por sala para cada curso avaliado de um total de 35 itens considerados.	74
4.2	Número de acertos, de respostas ausentes, total de respostas e percentual de acerto para cada item considerado de 277 respondentes.	75
4.3	Estimativas para os parâmetros a_i de discriminação, b_i de dificuldade e c_i de acerto ao acaso e seus respectivos intervalos de credibilidade (IC) de 95%.	77
4.4	Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por um desvio padrão.	83
4.5	Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por meio desvio padrão.	84
4.6	Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por meio desvio padrão.	85
4.7	Número de grupos selecionados, intervalo da escala para cada grupo, porcentagem de alunos em cada grupo e o número de itens âncora posicionados.	87
4.8	Número de grupos selecionados, intervalo da escala para cada grupo, porcentagem de alunos em cada grupo e o número de itens âncora posicionados.	87
4.9	Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 1.	89
4.10	Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 2.	90
4.11	Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 3.	91
4.12	Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 4.	92
4.13	Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 5.	93
4.14	Nome atribuído para cada grupo de acordo com a divisão da escala.	94
C.1	Número de acertos e proficiência por indivíduo na escala (250,50).	115
D.1	Gabarito, percentuais de resposta e parâmetros do item 2.	119
D.2	Gabarito, percentuais de resposta e parâmetros do item 3.	119
D.3	Gabarito, percentuais de resposta e parâmetros do item 4.	120
D.4	Gabarito, percentuais de resposta e parâmetros do item 5.	120
D.5	Gabarito, percentuais de resposta e parâmetros do item 6.	121

D.6 Gabarito, percentuais de resposta e parâmetros do item 7.	121
D.7 Gabarito, percentuais de resposta e parâmetros do item 8.	122
D.8 Gabarito, percentuais de resposta e parâmetros do item 9.	122
D.9 Gabarito, percentuais de resposta e parâmetros do item 10.	123
D.10 Gabarito, percentuais de resposta e parâmetros do item 11.	123
D.11 Gabarito, percentuais de resposta e parâmetros do item 12.	124
D.12 Gabarito, percentuais de resposta e parâmetros do item 13.	124
D.13 Gabarito, percentuais de resposta e parâmetros do item 14.	125
D.14 Gabarito, percentuais de resposta e parâmetros do item 15.	125
D.15 Gabarito, percentuais de resposta e parâmetros do item 16.	126
D.16 Gabarito, percentuais de resposta e parâmetros do item 17.	126
D.17 Gabarito, percentuais de resposta e parâmetros do item 18.	127
D.18 Gabarito, percentuais de resposta e parâmetros do item 19.	127
D.19 Gabarito, percentuais de resposta e parâmetros do item 20.	128
D.20 Gabarito, percentuais de resposta e parâmetros do item 21.	128
D.21 Gabarito, percentuais de resposta e parâmetros do item 22.	129
D.22 Gabarito, percentuais de resposta e parâmetros do item 23.	129
D.23 Gabarito, percentuais de resposta e parâmetros do item 24.	130
D.24 Gabarito, percentuais de resposta e parâmetros do item 25.	130
D.25 Gabarito, percentuais de resposta e parâmetros do item 26.	131
D.26 Gabarito, percentuais de resposta e parâmetros do item 27.	131
D.27 Gabarito, percentuais de resposta e parâmetros do item 28.	132
D.28 Gabarito, percentuais de resposta e parâmetros do item 29.	132
D.29 Gabarito, percentuais de resposta e parâmetros do item 31.	133
D.30 Gabarito, percentuais de resposta e parâmetros do item 32.	133
D.31 Gabarito, percentuais de resposta e parâmetros do item 33.	134
D.32 Gabarito, percentuais de resposta e parâmetros do item 34.	134
D.33 Gabarito, percentuais de resposta e parâmetros do item 35.	135
D.34 Gabarito, percentuais de resposta e parâmetros do item 36.	135

Lista de Siglas

- ANA: Avaliação Nacional da Alfabetização.
- ANAEB: Avaliação Nacional da Educação Básica.
- ANRESC: Avaliação Nacional do Rendimento Escolar.
- BIB: Blocos Incompletos Balanceados.
- BUGS: *Bayesian inference Using Gibbs Sampling*.
- CCI: Curva Característica do Item.
- EAP: Expected a Posteriori.
- EF: Ensino Fundamental.
- EM: Ensino Médio.
- EMV: Estimador de Máxima Verossimilhança.
- ENEM: Exame Nacional do Ensino Médio.
- IDEB: Índice de Desenvolvimento da Educação Básica.
- INEP: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
- MCMC: Métodos de Monte Carlo via Cadeias de Markov.
- MEC: Ministério da Educação.
- ML1: Modelo Logístico de 1 Parâmetro.
- ML2: Modelo Logístico de 2 Parâmetros.
- ML3: Modelo Logístico de 3 parâmetros.
- MV: Máxima Verossimilhança.
- MVC: Máxima Verossimilhança Conjunta.
- MVM: Máxima Verossimilhança Marginal.
- PDE: Plano de Desenvolvimento da Educação.
- SAEB: Sistema de Avaliação da Educação Básica.
- SEE/SP: Secretaria da Educação do Estado de São Paulo.
- SARESP: Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo.
- TCT: Teoria Clássica de Testes.
- TRI: Teoria da Resposta ao Item.

Sumário

Resumo	7
Abstract	9
Lista de Figuras	11
Lista de Quadros	13
Lista de Tabelas	15
Lista de Siglas	17
Capítulos	
1 Introdução	21
1.1 Construção de Escalas para Avaliação de Proficiência com base na Teoria da Resposta de Item	22
1.2 Teoria da Resposta ao Item	22
1.3 Construção e Interpretação da Escala para a Avaliação da Proficiência em Conteúdos Matemáticos Básicos	23
1.4 Apresentação da Dissertação	24
2 Construção de Escalas para Avaliação de Proficiências	25
2.1 Sistemas Brasileiros de Avaliação em Larga Escala	25
2.1.1 A Organização do SAEB	26
2.1.2 A Organização do ENEM	28
2.1.3 A Organização do SARESP	29
2.2 A Escala de Proficiência da TRI	30
2.2.1 Construção das Escalas de Proficiência do SAEB	32
2.2.2 Construção da Escala de Proficiência do ENEM	34
2.2.3 Construção das Escalas de Proficiência do SARESP	35
2.3 Proposta para Construção da Escala a partir de sua Segmentação	36
3 Teoria da Resposta ao Item: Conceitos, Definições e Estimação de Parâmetros	37
3.1 Modelos Matemáticos	38
3.2 O Modelo Logístico Unidimensional de três Parâmetros (ML3)	39
3.2.1 Curva Característica do Item - CCI: Interpretação e Representação Gráfica	40
3.2.2 Função de Informação do Item	43
3.2.3 Função de Informação do Teste	44

3.2.4	Suposições do Modelo	44
3.3	Estimação dos Parâmetros	45
3.3.1	Uma Breve Descrição dos Métodos Propostos na Literatura	45
3.3.2	Estimação Bayesiana	47
3.3.3	Descrição da Função de Verossimilhança	48
3.3.4	Descrição da Distribuição a Priori	48
3.3.5	Descrição da Distribuição a Posteriori	50
3.3.6	Descrição das Distribuições Condicionais Completas	51
3.3.7	Estimação das Proficiências	52
4	Construção e Interpretação da Escala de Proficiência em Conteúdos Matemáticos	53
4.1	Matriz de Referência	54
4.2	Elaboração do Instrumento de Avaliação	54
4.2.1	Tema I: Espaço e Forma	56
4.2.2	Tema II. Grandezas e Medidas	57
4.2.3	Tema III. Números e Operações/Álgebra e Funções	59
4.3	Aplicação da Avaliação	72
4.4	Discussão dos Resultados	73
4.4.1	Análise Descritiva dos Resultados	74
4.4.2	Análise dos Resultados Obtidos com Base na Teoria da Resposta ao Item	76
4.4.3	Posicionamentos de Itens na Escala	82
4.5	Construção e Interpretação da Escala de Proficiência em Conteúdos Matemáticos Básicos	86
4.5.1	Construção da Escala	86
4.5.2	Interpretação da Escala	94
5	Considerações Finais e Perspectivas Futuras	97
	Referências	99
	Apêndices	
A	Instrumento de Avaliação	103
B	Termo de Consentimento Livre e Esclarecido	113
C	Proficiência dos Indivíduos em função do Número de Acertos	115
D	Curva Característica e Curva de Informação de cada Item	119
	Anexos	
1	Cartão Resposta	137
2	Matriz de Referência	139

Introdução

Tradicionalmente, quando se deseja medir a proficiência de um indivíduo em determinada área do conhecimento, aplica-se uma prova (teste) com um determinado número de itens (questões) e com base no número de acertos determina-se seu escore. A pontuação obtida segue os princípios da Teoria Clássica de Testes (TCT). Neste contexto, a proficiência do indivíduo dependerá fortemente do teste ao qual foi submetido e comparações entre indivíduos submetidos a provas diferentes são difíceis de realizar.

Com o intuito de superar essas dificuldades e permitir uma medida mais adequada da proficiência de um indivíduo, entra em cena a Teoria da Resposta ao Item (TRI), que permite mensurar características que não podem ser medidas diretamente (traço latente) por meio de um conjunto de respostas a itens de um instrumento de avaliação. Desta forma, o elemento principal para esse tipo de análise é o item e não o instrumento.

Embora a TRI tenha surgido inicialmente dentro da Teoria Psicométrica em avaliações psicológicas, atualmente, é amplamente utilizada em avaliações educacionais. Com base na TRI, pode-se estabelecer escalas de proficiência interpretáveis que possibilitam, por exemplo, a comparação entre indivíduos e o acompanhamento da evolução dos sistemas de ensino ao longo dos anos.

Alguns exemplos de avaliação, utilizando a TRI, no âmbito internacional são: o exame de proficiência em língua inglesa - *Test of English as a Foreign Language* (TOEFL); a Avaliação Nacional do Progresso Educacional - *National Assessment of Educational Progress* (NAEP); o Programa Internacional de Avaliação de Estudantes - *Programme for International Student Assessment* (PISA) e o *Scholastic Aptitude Test* ou *Scholastic Assessment Test* (SAT), que consiste em um exame educacional padronizado, utilizado como critério para admissão nas universidades norte-americanas, semelhante ao ENEM no Brasil. No âmbito nacional, o Sistema de Avaliação da Educação Básica (SAEB), o Exame Nacional do Ensino Médio (ENEM) e o Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP), são exemplos de sistemas de avaliação educacional que utilizam a TRI.

Mais especificamente, a TRI consiste em uma família de modelos matemáticos que relaciona variáveis observáveis (as respostas dadas aos itens de uma prova, por exemplo) e variáveis que não podem ser observadas diretamente (a proficiência de um indivíduo em matemática básica, por exemplo). As etapas para a aplicação dos modelos TRI envolvem a construção de uma matriz de referência para o traço latente (proficiência), a elaboração de itens para a construção da prova, a aplicação da prova e a leitura dos dados, a especificação do modelo, a estimação dos parâmetros dos itens e da proficiência dos respondentes e, finalmente, a construção e a interpretação da escala.

Considerando a importância do tema, os objetivos deste trabalho consistem em um estudo sobre construções de escalas baseadas na TRI com ênfase na área educacional e a partir disto, a construção e interpretação de uma escala de proficiência em conteúdos matemáticos básicos.

Neste capítulo apresenta-se uma síntese do conteúdo deste trabalho e a motivação para o seu desenvolvimento.

1.1 Construção de Escalas para Avaliação de Proficiência com base na Teoria da Resposta de Item

Como já mencionado, as etapas para a aplicação dos modelos TRI envolvem a construção de uma matriz de referência para o traço latente em questão, a elaboração de itens para a construção da prova, a aplicação da prova, a leitura dos dados, a especificação do modelo a ser utilizado, a estimação dos parâmetros dos itens e da proficiência dos respondentes e, finalmente, a construção e interpretação da escala.

As escalas de conhecimento (ou proficiência) são escalas que possibilitam a interpretação pedagógica dos valores das proficiências estimadas. Estas escalas de conhecimento são definidas a partir de níveis selecionados para serem interpretados pedagogicamente, de acordo com os itens que forem posicionados, seguindo algum critério.

Alguns critérios propostos na literatura e discutidos neste trabalho envolvem o proposto por Beaton e Allen (1992), os critérios de posicionamento utilizados pelos grandes sistemas de avaliação educacional, a saber o SAEB, o ENEM e o SARESP e o critério adotado nesta pesquisa, que foi o de segmentar a escala de proficiência proposta em faixas de proficiência a partir da análise de agrupamentos hierárquicos.

O SAEB, criado em 1990, é composto por três avaliações externas em larga escala. Passou a utilizar a TRI a partir do ano de 1995, permitindo a elaboração de escalas de proficiências comuns entre anos e comparáveis com anos anteriores.

O ENEM é aplicado anualmente desde 1998, pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e Ministério da Educação (MEC). Sua finalidade é avaliar o desempenho escolar ao final do Ensino Médio. Começou-se a utilizar a TRI a partir de 2009, quando foram estabelecidas as matrizes de referência para cada uma de suas áreas do conhecimento avaliadas.

O SARESP é aplicado anualmente pela Secretaria do Estado de São Paulo (SEE/SP) desde 1996, com o objetivo de produzir um diagnóstico da situação da escolaridade básica paulista. Utiliza a TRI em suas avaliações desde sua implantação. A partir da criação de escalas de proficiências interpretáveis, os resultados obtidos ao longo dos anos nas séries/anos consecutivos são comparados.

1.2 Teoria da Resposta ao Item

Segundo Andrade, Tavares e Valle (2000), os primeiros modelos da TRI surgiram na década de 50. Estes modelos consideravam apenas uma única proficiência de um único grupo (ou população) e itens corrigidos de forma dicotômica (certo ou errado). Inicialmente, foram especificados na forma de uma função ogiva normal, e mais tarde descritos na forma da função logística. Atualmente, os modelos propostos na literatura consideram itens dicotômicos ou não, uma ou mais populações e um ou mais traços latentes.

Nas avaliações educacionais, os testes (provas) mais comuns são os compostos por itens de múltipla escolha, corrigidos como certo ou errado, ou seja, de maneira dicotômica. A TRI busca inferir a proficiência de um indivíduo em um determinado contexto. Para isso, utiliza um modelo matemático adequado, que expressa a probabilidade de resposta correta a um item em função do traço latente do indivíduo e dos parâmetros do item. No âmbito educacional, o modelo matemático comumente utilizado e adotado nesta pesquisa é o logístico unidimensional de três parâmetros, que leva em consideração a discriminação (parâmetro a), a dificuldade (parâmetro b) e a probabilidade de acerto ao acaso (parâmetro c) do item.

Quando esses parâmetros são desconhecidos, tanto dos itens quanto das proficiências dos indivíduos, métodos de estimação precisam ser utilizados para fazer afirmações sobre os seus valores. Graças ao avanço tecnológico, ao longo dos anos, foram desenvolvidos software que facilitaram em muito a aplicação da TRI, pois devido à complexidade desses modelos se faz necessário o uso de métodos numéricos para o processo de inferência sobre os parâmetros dos itens e da proficiência dos indivíduos.

O processo de inferência sobre os parâmetros pode ser feito de acordo com a abordagem clássica ou bayesiana. Na abordagem clássica, em geral, a estimação dos parâmetros de interesse é feita pelo Método da Máxima Verossimilhança. Neste trabalho optou-se pela abordagem bayesiana, utilizando-se o amostrador de Gibbs, algoritmo da classe dos Métodos de Monte Carlo via Cadeia de Markov (MCMC), já implementado no software OpenBUGS (Bayesian Inference Using Gibbs Sampler) (LUNN et al., 2009). A ideia básica do MCMC é gerar amostras da distribuição a posteriori, a partir de distribuições condicionais completas, que formam o núcleo de uma cadeia de Markov. Sob condições gerais de regularidade a distribuição da cadeia gerada pelo amostrador de Gibbs, converge para a distribuição de equilíbrio, neste caso a posteriori conjunta.

Para a comparação dos resultados clássicos e bayesianos utilizou-se o software BILOG-MG (inferência clássica), este também utilizado pelo INEP.

1.3 Construção e Interpretação da Escala para a Avaliação da Proficiência em Conteúdos Matemáticos Básicos

A edição 2016 do SARESP aferiu a proficiência média em Matemática dos estudantes da 3ª Série do Ensino Médio como 278,1. Essa proficiência média está associada ao nível de proficiência Básico. No entanto, deve-se ressaltar que um grande número de alunos tem sua proficiência no nível abaixo do básico ou básico como pode ser visto na Tabela 1.1. Esses resultados mostram que a maior parte dos alunos saem do Ensino Médio com defasagem nos conteúdos matemáticos básicos (SÃO PAULO, 2017b).

Tabela 1.1: Nível de Proficiência em Matemática dos alunos do 3º ano do Ensino Médio (SARESP 2016).

Nível de Proficiência	%
Abaixo do Básico	47,6
Básico	47,3
Adequado	4,8
Avançado	0,3

O relatório do vestibular UNESP de 2016, emitido pela Fundação VUNESP, informa através do questionário socioeconômico que 47,1% dos alunos ingressantes nos cursos das áreas de exatas cursaram o ensino médio todo ou em maior parte em escolas públicas.

Esses dados e os altos índices de retenção e evasão em disciplinas do núcleo básico dos cursos de graduação na área de exatas, motivaram a construção e a interpretação de uma escala de proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo e similares. Este trabalho, dá continuidade à pesquisa desenvolvida por Rossi (2015), que construiu a matriz de referência utilizada nesta pesquisa e aplicou uma avaliação composta por 32 itens com o objetivo de testá-los.

A matriz de referência está dividida em três grupos de conteúdos (Espaço e Forma, Grandezas e Medidas e Números e Operações/Álgebra e Funções) e estes divididos por competências. Cada competência descreve uma habilidade em conteúdos matemáticos básicos que se deseja inferir. Neste trabalho foram propostos 36 itens, um item para cada uma das habilidades descritas na matriz de referência. Todos os itens são de múltipla escolha com cinco alternativas cada um, sendo apenas uma a correta. Antes das aplicações das provas, foram feitas reuniões com os professores do Departamento de Matemática e Computação da FCT/UNESP que analisaram as questões propostas, a fim de corrigir possíveis erros e avaliar se estas estavam adequadas para inferir as habilidades esperadas.

A prova foi aplicada aos alunos matriculados no primeiro semestre do primeiro ano dos cursos da área de exatas da FCT/UNESP, no ano de 2017, totalizando 8 cursos, nos dias e horários de aulas disponibilizados pelos professores responsáveis pelas disciplinas de Cálculo diferencial e Integral I e Matemática Elementar I, durante os meses de agosto e setembro 2017.

Conforme já mencionado, a construção da escala se dá após a estimação dos parâmetros dos itens e das proficiências. Para isso, foram apresentadas formas de posicionamento de itens e escolha de níveis, denominados âncora. Após a análise destas formas, optou-se por dividir a escala em faixas de proficiência, a partir da análise de agrupamentos hierárquicos, possibilitando a interpretação da escala em toda a sua amplitude.

1.4 Apresentação da Dissertação

O objetivo geral deste trabalho consistiu em um estudo sobre construções de escalas baseadas na TRI com ênfase na área educacional. Como objetivo específico a construção e interpretação de uma escala de proficiência em conteúdos matemáticos básicos.

Para atingir os objetivos propostos, apresenta-se no Capítulo 2 uma descrição de como estão organizados e a construção das escalas de proficiências de três grandes sistemas brasileiros de avaliação educacional. Também é apresentado neste capítulo a metodologia adotada para interpretação da escala proposta nesta pesquisa. No Capítulo 3 apresenta-se o modelo logístico unidimensional de três parâmetros e o método de estimação, sob o enfoque bayesiano, adotados para este trabalho. Também é apresentada uma síntese sobre modelos propostos na literatura e sobre outros métodos de estimação.

No Capítulo 4 apresenta-se a construção da escala e interpretação da mesma, envolvendo a apresentação da matriz de referência, a elaboração e aplicação da prova e a discussão dos resultados. E por fim, no Capítulo 5 são apresentadas as considerações finais e perspectivas futuras.

Construção de Escalas para Avaliação de Proficiências

Com base na TRI pode-se estabelecer uma escala métrica padrão para mensurar o conhecimento, permitindo que a comparação de resultados entre avaliações se torne possível. Sabe-se que os conhecimentos adquiridos por um estudante não podem ser mensurados diretamente, mais pode-se utilizar instrumentos de medida que busquem mensurá-los indiretamente. Sendo estes instrumentos de medida do conhecimento (traço latente), as provas aplicadas nas avaliações educacionais.

O uso da TRI em avaliações educacionais no Brasil iniciou-se em 1995 com o SAEB. Desde então, a TRI vem sendo amplamente utilizada na área educacional. Embora, seja vasta sua disseminação em abordagens educacionais, pode-se encontrar na literatura aplicações da TRI nas mais diversas áreas tais como no grau de satisfação, avaliação de intangíveis nas empresas, gestão organizacional, gestão escolar, avaliação de intenções comportamentais, avaliação de atitudes, orientação profissional, avaliação de qualidade de vida, avaliação de sintomas depressivos ou ansiedade, gestão de qualidade, sociologia, nível socioeconômico, raciocínio diagnóstico, usabilidade de sites, avaliação de emissão de poluentes, avaliação de itens constrangedores, avaliações psicológicas diversas, inteligência emocional, percepção emocional, na área da saúde, na medicina, na genética, na epidemiologia, na odontologia, na avaliação de raciocínio verbal, no marketing, na gestão do conhecimento, na avaliação empresarial e também em testes adaptativos informatizados. Referências bibliográficas nestas áreas podem ser consultadas no estudo publicado por (MOREIRA JUNIOR, 2010).

Neste capítulo será apresentada a escala de proficiência da TRI e sua aplicação em três grandes sistemas de avaliação brasileiros. Sendo estes, o SAEB, o ENEM e o SARESP e a proposta utilizada neste trabalho para a construção da escala. Desta forma, na Seção 2.1 apresenta-se a organização destes três sistemas de avaliação, na Seção 2.2 são apresentadas as escalas da TRI e a construção das escalas de proficiência destes sistemas educacionais e na Seção 2.3 é apresentada a proposta de segmentação da escala utilizada nesta pesquisa.

2.1 Sistemas Brasileiros de Avaliação em Larga Escala

O SAEB, o ENEM e o SARESP são os principais sistemas brasileiros de avaliação em larga escala que fazem uso da TRI. Na sequência descreve-se como cada um destes sistemas está organizado.

2.1.1 A Organização do SAEB

O SAEB foi criado em 1990. Seu principal objetivo é realizar um diagnóstico da educação básica brasileira. Produzindo informações que subsidiem a formulação, reformulação e o monitoramento das políticas públicas nas esferas municipal, estadual e federal, tendo em vista contribuir para a melhoria da qualidade, equidade e eficiência do ensino. Além disso, fornece dados e indicadores sobre fatores de influência do desempenho dos alunos, nas áreas e anos avaliados (BRASIL, 2017a).

De acordo com Brasil (2017a) o SAEB foi reestruturado em 2005 e passou a ser composto por duas avaliações: a Avaliação Nacional da Educação Básica (ANEB), que manteve as características, os objetivos e os procedimentos da avaliação efetuada até então, avaliando escolas públicas e particulares e a Avaliação Nacional do Rendimento Escolar (ANRESC), conhecida como Prova Brasil, criada com o objetivo de avaliar a qualidade do ensino ministrado somente nas escolas das redes públicas. Posteriormente, em 2013, foi incorporada ao SAEB a Avaliação Nacional da Alfabetização (ANA), com o objetivo de avaliar os níveis de alfabetização e letramento em Língua Portuguesa, alfabetização Matemática e condições de oferta do Ciclo de Alfabetização das redes públicas, aplicadas somente no final do ciclo de alfabetização. As disciplinas abordadas atualmente nos três tipos de avaliação são as disciplinas de Língua Portuguesa (foco em leitura) e de Matemática (foco na resolução de problemas). A partir de informações obtidas em Brasil (2017a), descreve-se a seguir um pouco sobre cada uma dessas avaliações.

A ANRESC ou Prova Brasil é uma avaliação bianual, com abrangência censitária. Envolve alunos do 5º e 9º ano do Ensino Fundamental (EF) das escolas públicas com no mínimo 20 alunos matriculados em cada um dos anos. Tem por objetivo principal, mensurar a qualidade do ensino ministrado nas escolas das redes públicas, fornecendo resultados para cada unidade escolar participante, bem como para as redes de ensino em geral. Também, apresenta, indicadores contextuais sobre as condições extra e intraescolares em que ocorre o trabalho da escola.

A ANEB ou Prova do SAEB utiliza os mesmos instrumentos da Prova Brasil e é aplicada com a mesma periodicidade. Diferencia-se por abranger, de forma amostral, escolas e alunos das redes públicas e privadas do País, em áreas urbanas e rurais, matriculados no 5º e 9º ano do EF e no 3º ano do Ensino Médio (EM) regular. Essa prova mantém as características, os objetivos e os procedimentos efetuados até 2005 pelo SAEB. Tem como foco, avaliar a qualidade, a equidade e a eficiência da educação básica brasileira, sendo os resultados apresentados por regiões geográficas do País e unidades da federação. Para os 5º e 9º anos do EF de escolas públicas, urbanas ou rurais, é aplicada em turmas com 10 a 19 alunos e nos 3º anos do EM nas escolas públicas, urbanas e rurais, nas turmas com 10 alunos ou mais. Nos 5º e 9º anos do EF e 3º anos do EM das escolas privadas, urbanas e rurais, em turmas com 10 alunos ou mais, abrangendo assim, escolas e alunos, que não atendem aos critérios de participação da Prova Brasil.

Para a aplicação da Prova Brasil e da ANEB foram construídas matrizes de referência para cada uma das áreas do conhecimento abordadas (Língua Portuguesa e Matemática), estruturadas de acordo com as séries avaliadas. Para cada uma delas, são definidos o conjunto de conteúdos (tópicos ou temas) e habilidades a serem avaliadas representando o que se espera dos alunos ao final dos anos avaliados. Para a construção dessas matrizes de referência o INEP baseou-se em Parâmetros Curriculares Nacionais, consulta nacional aos currículos propostos pelas Secretarias Estaduais de Educação e por algumas redes municipais. Ainda, foram consultados professores regentes das redes e examinados os livros didáticos mais utilizados para os anos avaliados. As matrizes de referência, como o próprio nome sugere, são utilizadas como referência para a elaboração dos itens (questões)

da prova. Desta forma, não devem ser confundidas com as propostas curriculares das redes ou das escolas, pois não englobam todo o currículo escolar.

Com base na matriz de referência de cada área do conhecimento e ano avaliado, são elaborados os itens que compõem os testes. Para montar os cadernos de prova, o INEP utiliza uma metodologia conhecida como Blocos Incompletos Balanceados (BIB), com o objetivo de permitir que um grande número de itens seja aplicado ao conjunto de alunos avaliados, sem que estes precisem responder a todas as questões que cobrem a matriz de referência. Além disso, o BIB permite uniformizar a exposição dos itens, blocos e cadernos aos alunos respondentes.

Para a avaliação do 5º ano por exemplo, o processo de confecção dos cadernos de prova para cada uma das áreas do conhecimento é realizado do seguinte modo: são montados 7 blocos com 11 itens cada, totalizando 77 itens de múltipla escolha para cada área. Feito isto, cada caderno é montado agrupando-se dois blocos de cada área do conhecimento. De acordo com a metodologia utilizada, obtém-se 21 cadernos de prova diferentes. Cada aluno avaliado responde somente a um caderno no dia da avaliação. Assim, o caderno contém 22 questões de Língua Portuguesa e 22 questões de Matemática. Para o 9º ano, o caderno é montado seguindo a mesma metodologia, porém, cada bloco contém 13 itens, totalizando 91 itens de múltipla escolha para cada área. Desta forma, o caderno de prova respondido pelos alunos do 9º ano é composto por 26 itens de Língua Portuguesa e 26 itens de Matemática.

As provas da ANEB e da Prova Brasil são acompanhadas de questionários que são respondidos pelos alunos, professores dos anos avaliados e diretores das unidades escolares. Os aplicadores da prova também respondem a um questionário da escola. Este levantamento de dados é fundamental para que a análise da Prova Brasil leve em consideração o contexto dos alunos e da escola.

O INEP, realizou em novembro de 2015, a 13ª edição do SAEB. De acordo com o Censo da Educação Básica 2015, esta edição, contou com a participação de todas as escolas públicas brasileiras com, pelo menos, 20 alunos matriculados no 5º ou 9º anos do EF. Participaram também, uma amostra de escolas privadas com 10 ou mais estudantes matriculados no 5º ou 9º anos do EF ou no 3º ano do EM, assim como uma amostra de escolas públicas municipais e estaduais com 10 a 19 alunos matriculados no 5º ou 9º anos do EF e uma amostra de escolas públicas municipais e estaduais com 10 ou mais alunos matriculados no 3º ano do EM, totalizando 3.986.190 estudantes participantes.

Para a ANA, avaliação censitária anual, aplicada para o 3º ano do EF das escolas públicas, foram realizadas apenas três edições até o momento. A primeira ocorreu no ano de sua criação em 2013, a segunda em 2014 e a terceira em 2016. Nesta última edição, participaram da ANA, todas as escolas públicas, urbanas e rurais que, até o dia 31 de agosto do mesmo ano, haviam registrado no sistema Educacenso pelo menos 10 estudantes matriculados em turmas regulares do 3º ano do EF, sendo esta a última etapa do ciclo de alfabetização. Segundo Brasil (2017c), nesta edição os testes foram aplicados para 2,5 milhões de estudantes referentes a 50 mil escolas e 100 mil turmas.

O Educacenso é uma ferramenta que permite o detalhamento do sistema educacional brasileiro. É possível obter dados individualizados de cada estudante, professor, turma e escola do País, tanto das redes públicas (federal, estaduais e municipais) quanto da rede privada. A partir dos dados do Educacenso é feito o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB) e o planejamento da distribuição de recursos para alimentação, transporte escolar e livros didáticos, entre outros (BRASIL, 2017b).

As matrizes de referência da ANA, assim como, da ANEB e da Prova Brasil, foram elaboradas a partir de contribuições de diferentes pesquisadores e especialistas de várias universidades, no campo da alfabetização e letramento. Também colaboraram para a ela-

boração das matrizes, representantes de diversas instituições do governo e da sociedade civil como o MEC, o Conselho Nacional de Educação (CNE), a União Nacional dos Dirigentes Municipais de Educação (UNDIME), a Associação Nacional de Pós-Graduação e Pesquisa em Educação (ANPED), a União Nacional dos Conselhos Municipais de Educação (UNCME), o Conselho Nacional de Secretários de Educação (CONSED), as Secretarias de Educação, dentre outras.

Os alunos que participam da ANA respondem a um teste composto por 20 itens de língua Portuguesa, sendo 17 itens objetivos de múltipla escolha e 3 itens de produção escrita e 20 itens de Matemática, sendo estes todos de múltipla escolha. Para a coleta de informações a respeito das condições de oferta, são aplicados questionários direcionados aos professores e gestores das instituições de ensino que atendem ao ciclo de alfabetização. Os questionários tem como foco verificar informações sobre as condições de infraestrutura; formação de professores; gestão da unidade escolar e organização do trabalho pedagógico.

2.1.2 A Organização do ENEM

O ENEM tem por finalidade principal a avaliação do desempenho escolar e acadêmico ao fim do Ensino Médio. O exame é realizado anualmente desde de 1998 pelo INEP e MEC. Qualquer pessoa pode fazer o exame, contudo, os participantes menores de 18 anos que não concluirão o EM no mesmo ano letivo de sua participação no exame, só poderão usar os resultados para uma autoavaliação de conhecimentos.

A avaliação do ENEM é composta por questões objetivas de múltipla escolha e uma redação. A partir de 2009, as provas objetivas do ENEM passaram a ser estruturadas com base em matrizes de referência. Assim, foram construídas matrizes de referência para quatro áreas de conhecimento sendo elas:

- Linguagens, códigos e suas tecnologias, que abrange o conteúdo de Língua Portuguesa (Gramática e Interpretação de Texto), Língua Estrangeira Moderna, Literatura, Artes, Educação Física e Tecnologias da Informação;
- Matemática e suas tecnologias;
- Ciências da Natureza e suas tecnologias, que abrange os conteúdos de Química, Física e Biologia;
- Ciências Humanas e suas tecnologias, que abrange os conteúdos de Geografia, História, Filosofia e Sociologia.

As questões que compõem os cadernos de prova são construídas com base nas matrizes de referência para cada área do conhecimento. São 45 questões de múltipla escolha para cada uma dessas áreas, divididas em dois dias de aplicação. A partir de 2017 as provas foram organizadas da seguinte forma: No primeiro dia são 45 questões da área de Linguagens, códigos e suas tecnologias e 45 questões da área de Ciências Humanas e suas tecnologias, totalizando 90 questões de múltipla escolha mais uma proposta de redação. No segundo dia são aplicadas 45 questões para área de Matemática e suas tecnologias e 45 questões da área Ciências da Natureza e suas Tecnologias totalizando 90 questões de múltipla escolha.

As redações são avaliadas de acordo com 5 competências sendo que, dois corretores independentes atribuem notas de zero (0) a duzentos (200) pontos para cada uma das competências, sendo elas:

1. Demonstrar domínio da norma padrão da língua escrita.

2. Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento, para desenvolver o tema dentro dos limites estruturais do texto dissertativo-argumentativo.
3. Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista.
4. Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação.
5. Elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos.

As informações obtidas a partir dos resultados do ENEM são utilizadas para permitir a autoavaliação do participante, para continuidade de sua formação e inserção no mercado de trabalho, para criar uma referência nacional para o aperfeiçoamento dos currículos do Ensino Médio, para acesso a educação superior, como mecanismo único, alternativo ou complementar, para acesso a programas governamentais de financiamento ou o apoio ao estudante da educação superior, acesso a Instituições de Ensino Superior Portuguesas, para ingresso em diferentes setores do mundo do trabalho e para desenvolver estudos e indicadores sobre a educação brasileira. As informações fornecidas aqui, podem ser obtidas em (BRASIL, 2017d)

2.1.3 A Organização do SARESP

Segundo São Paulo (2017), o SARESP, aplicado pela SEE/SP, tem como objetivo produzir um diagnóstico da situação da escolaridade básica paulista, a fim de contribuir para melhoria da qualidade da educação. São avaliados atualmente pelo SARESP os alunos de 3º, 4º, 5º, 7º e 9º anos do EF e 3º ano do EM da rede estadual. Os resultados obtidos são utilizados para orientar ações a respeito da educação e também integrar o cálculo do Índice de Desenvolvimento da Educação do Estado de São Paulo (IDESP). O IDESP tem o papel de dialogar com a escola, fornecendo dados sobre sua qualidade, aponta os pontos em que precisa melhorar e sinaliza sua evolução ano a ano.

A avaliação do SARESP é censitária e vem sendo aplicada anualmente desde 1996. O sistema passou por algumas mudanças no ano de 2007, com isso, a partir de 2008, a avaliação passou a ser aplicada para todas as áreas curriculares, alternado ano a ano a periodicidade das áreas avaliadas. Anualmente são avaliadas as disciplinas de Língua Portuguesa, redação e Matemática e, anual e alternadamente as áreas de Ciências da Natureza, contemplando as disciplinas de Ciências, Física, Química e Biologia e Ciências Humanas, contemplando as disciplinas de História e Geografia. Também foi feita a adequação da matriz de referência do 5º e 9º ano do EF e 3º ano do EM às matrizes de referência do SAEB/Prova Brasil.

Para a construção das matrizes de referência do SARESP, assim como nas avaliações citadas anteriormente, contou-se com a colaboração de vários especialistas. Uma primeira proposta foi organizada por especialistas em avaliação, tendo por base a Proposta Curricular do Estado de São Paulo. Considerou-se também, documentos que cercam as avaliações nacionais e internacionais. Esta Primeira versão das matrizes foram apresentadas aos autores da proposta curricular, que contribuíram com uma leitura crítica do documento, especialistas da Coordenadoria de Estudos e Normas Pedagógicas (CENP) da SEE/SP também realizaram a leitura do documento, fornecendo sugestões e ajustes. Além disso, as matrizes foram discutidas em reuniões técnicas (oficinas) com professores

coordenadores de oficinas pedagógicas das áreas envolvidas na avaliação. Deste cuidadoso trabalho realizado por diferentes equipes, resultou as matrizes de referência do SARESP.

Para cada um dos anos/séries avaliados foram construídas matrizes de referência em cada uma das áreas do conhecimento avaliadas. Para cada um deles, são definidos o conjunto de conteúdos (tópicos ou temas) e as habilidades a serem examinadas em cada uma das áreas. Essas habilidades foram classificadas em três grupos. O primeiro grupo é o grupo das competências para observar, o segundo para realizar e o terceiro para compreender. Para a redação, são levadas em consideração a avaliação de cinco competências, sendo elas tema, gênero, coesão/coerência, registro e proposição.

A edição de 2016 do SARESP contou com a participação de 956.820 alunos distribuídos em 5.105 escolas da rede Estadual, 73.894 alunos distribuídos em 514 escolas das redes municipais, 19.451 alunos de 200 Escolas Técnicas Estaduais (ETE) administradas pelo Centro Estadual de Educação Tecnológica Paula Souza e vinculados à Secretaria Estadual do Desenvolvimento do Estado de São Paulo e 1.869 alunos de 20 escolas particulares. A participação das escolas particulares se deu por adesão voluntária das mesmas.

No ano de 2016, assim como no ano anterior 2015, foram avaliadas somente as disciplinas de Língua Portuguesa e Matemática. A avaliação censitária englobou alunos do 3º, 5º e 9º anos do EF e 3º ano do EM. A avaliação dos alunos do 7º do EF da rede estadual da SEE/SP foi realizada por amostragem, com uma amostra correspondente a 10% da população em questão.

Para o 3º ano do EF o caderno de prova para a avaliação de Língua Portuguesa era composto por 8 questões de resposta aberta e 5 questões objetivas, de múltipla escolha. Para a avaliação de Matemática, os cadernos de prova eram compostos com 13 questões de resposta aberta e 5 questões objetivas. Essas provas foram corrigidas online por professores especialistas.

Os alunos do 5º, 7º e 9º anos do EF, e da 3ª série do EM foram avaliados por 104 questões objetivas de Língua Portuguesa e 104 questões objetivas de Matemática. Estas, foram divididas em 13 blocos contendo 8 questões cada um, para cada uma das áreas de conhecimento. Cada caderno de prova foi montado segundo a metodologia de Blocos Incompletos Balanceados, agrupando-se três blocos de cada área. Com a utilização da metodologia adotada, foram montados 26 cadernos diferentes contendo 24 itens de língua portuguesa e 24 itens de Matemática.

2.2 A Escala de Proficiência da TRI

Na TRI a proficiência de um indivíduo pode assumir teoricamente qualquer valor real entre $-\infty$ e $+\infty$, diferente do que acontece na medida escore de um teste com I questões corrigidas como certo ou errado, onde a proficiência assume valores de acordo com pontuações propostas inicialmente. Seguindo a metodologia da TRI, para a definição de uma escala é necessário estabelecer uma origem e uma unidade de medida, escolhidas de modo que representem, respectivamente, o valor médio e o desvio padrão dos indivíduos da população em estudo.

Na prática, frequentemente, se utiliza a escala $(0,1)$. Isto é, a escala com média igual a 0 e desvio padrão igual a 1. Nesta escala, os valores do parâmetro b variam (tipicamente) entre -3 e $+3$ e espera-se que o parâmetro a assumam valores maiores que 0, sendo considerados os valores de a maiores que 1 os mais apropriados segundo Andrade, Tavares e Valle (2000). No entanto, não faz a menor diferença que se estabeleça outros valores quaisquer de média e desvio padrão para a escala, o que importa de fato são as relações de ordem existentes entre esses pontos. Por exemplo, um indivíduo com proficiência 1,5 na escala $(0,1)$ está 1,5 desvios padrão acima da média. Este mesmo

indivíduo numa escala (500, 100) (escala utilizada pelo ENEM) terá proficiência 650, que continua sendo 1,5 desvios padrão acima da média.

A escala (0, 1), como citado no parágrafo anterior, na prática é a mais utilizada para construção de escalas de proficiência. Esta, por sua vez, não se apresenta de forma tão elegante, pois exibirá níveis de proficiência com valores negativos e também números decimais, o que pode causar certo desconforto na utilização da escala. Deste modo, após a estimação dos parâmetros dos itens (também conhecida como calibração dos itens) e das proficiências utilizando a escala (0, 1), é possível efetuar uma transformação para qualquer outra escala que o pesquisador desejar utilizar, sendo que este procedimento tem como único propósito, facilitar a construção e utilização da escala, trabalhando com números positivos e inteiros. Essa transformação é feita de maneira que

$$a(\theta - b) = (a/\sigma)[(\sigma \times \theta + \mu) - (\sigma \times b + \mu)]$$

onde μ e σ , representam a média e o desvio padrão respectivamente e $a(\theta - b)$ é a parte do modelo probabilístico proposto envolvida na transformação. Portanto, na escala transformada tem-se que

$$(i) \theta^* = \sigma \times \theta + \mu;$$

$$(ii) b^* = \sigma \times b + \mu;$$

$$(iii) a^* = a/\sigma;$$

$$(iv) P(U_i = 1 | \theta) = P(U_i = 1 | \theta^*)$$

U_i é a variável dicotômica (corrigida como certo ou errado) que assume valor 1 quando o indivíduo com proficiência θ responde corretamente ao item i e 0 caso contrário.

É importante ressaltar que independente da escala com a qual se está trabalhando, a probabilidade de um indivíduo responder corretamente a um certo item é sempre a mesma, ou seja, a proficiência de um indivíduo é invariante à escala de medidas.

Para que seja possível construir escalas de proficiência, é necessário que todos os parâmetros dos itens e que todas as proficiências dos indivíduos respondentes (tanto individuais como populacionais) de todos os grupos avaliados, estejam numa mesma escala de medida, ou seja, que todos os parâmetros envolvidos possam ser comparados.

As estimativas dos parâmetros dos itens e das proficiências segue uma natureza arbitrária. Devido a este fato, a menos que se faça uma ligação desses valores obtidos com os conteúdos envolvidos na prova, é possível somente comparar entre si as proficiências obtidas para os diferentes indivíduos. Porém, não é possível qualquer interpretação pedagógica. Sem essa ligação, pode-se dizer apenas que um indivíduo que possui proficiência 1,5 na escala (0, 1) deve possuir um conhecimento maior sobre o conteúdo avaliado do que um indivíduo que possui proficiência -1 , por exemplo, e que o indivíduo com proficiência 1,5 está a 1,5 desvios padrão acima da média de proficiência da população avaliada, ao mesmo tempo que o indivíduo com proficiência -1 está a 1 desvio padrão abaixo da média das proficiências dessa mesma população. Mas, não podemos afirmar nada a respeito do que o primeiro indivíduo sabe a mais do que o segundo (ANDRADE; TAVARES; VALLE, 2000).

Para tornar possível a interpretação pedagógica dos valores das proficiências, criou-se as escalas de conhecimento (ou escalas de proficiências). As escalas de conhecimento são definidas a partir de níveis, conhecido como níveis âncoras, que são definidos por itens

conhecidos como itens âncora. Os níveis âncora são pontos na escala selecionados pelo pesquisador (analista) para serem interpretados pedagogicamente. Já os itens âncora, podem ser definidos de acordo com algum critério como o da definição sugerida por Beaton e Allen (1992) e também apresentada por Andrade, Tavares e Valle (2000).

Definição: *Considere dois níveis âncora consecutivos Y e Z com $Y < Z$. Dizemos que um determinado item é âncora para o nível Z se, e somente se, as 3 condições abaixo forem satisfeitas simultaneamente*

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

$$(2) P(U = 1 | \theta = Y) < 0,50 \text{ e}$$

$$(3) P(U = 1 | \theta = Z) - P(U = 1 | \theta = Y) \geq 0,30.$$

Ou seja, para um item ser âncora de um determinado nível âncora da escala, ele precisa ser respondido corretamente por uma grande proporção de indivíduos ($\geq 65\%$) com este nível de proficiência Z considerado (satisfazendo a primeira condição). Deve ser respondido corretamente por uma proporção menor de indivíduos ($< 50\%$) com um nível de proficiência imediatamente inferior (satisfazendo a segunda condição). Além disso, a diferença entre a proporção de indivíduos que responderam corretamente esse item com esses níveis de proficiência deve ser maior que 30% (satisfazendo a terceira e última condição).

Previamente, não se pode ter certeza de quantos itens âncoras serão selecionados para cada um dos níveis âncora escolhidos e nem se serão selecionados itens âncora (no teste aplicado) para todos os níveis âncora fixados. Diante disto, é essencial que os níveis âncora sejam escolhidos de maneira que não fiquem muito próximos uns dos outros e também que sejam aplicados uma grande quantidade de itens, a fim de possibilitar a construção e interpretação da escala de proficiência.

2.2.1 Construção das Escalas de Proficiência do SAEB

A TRI começou a ser utilizada nas avaliações aplicadas pelo SAEB a partir do ano de 1995. O emprego da TRI permitiu a elaboração de escalas de proficiência comuns entre as séries e na mesma escala obtida em anos anteriores. Possibilitando assim, a comparação e acompanhamento da evolução do sistema de ensino ao longo dos anos.

Para que os alunos de todas as séries e de todos os anos sejam postos em uma mesma escala de proficiência, de maneira que seus desempenhos possam ser comparados, o SAEB utiliza itens comuns entre séries e anos avaliados viabilizando, assim, o um processo denominado equalização. O modelo utilizado para estimação dos parâmetros e das proficiências é o modelo Logístico Unidimensional de três parâmetros, o processo de estimação utilizado é a Estimação por Máxima Verossimilhança Marginal onde são atribuídas distribuições a priori para os parâmetros dos itens e proficiências. Os parâmetros dos itens são estimados de forma conjunta, para todos os itens de todas as séries de uma mesma disciplina com a utilização do programa BILOG-MG que, implementa uma extensão da TRI a grupos múltiplos de respondentes, onde as diferentes séries/anos são os grupos. Esta extensão da TRI de uma população (ou grupo) para várias populações (ou grupos) diferentes é feita atribuindo-se distribuições a priori distintas para cada grupo. Com o intuito de evitar a indeterminação do modelo, é atribuída uma distribuição a priori para um grupo tomado como grupo de referência. Desta maneira, é feita a estimação conjunta dos parâmetros das distribuições a priori dos outros grupos e dos parâmetros dos itens. Este procedimento permite que se faça simultaneamente a equalização entre séries e, quando

for o caso, a equalização entre anos. Para mais detalhes sobre a TRI aplicada a grupos múltiplos pode-se consultar (BOCK; ZIMOWSKI, 1997).

Klein (2009) traz uma descrição detalhada de como ocorreu o processo de construção da escala do SAEB nos anos de 1995, 1997, 1999 e 2001. A primeira escala comum entre séries e entre anos utilizada pelo SAEB foi baseada na avaliação aplicada em 1997, onde a média e o desvio padrão da distribuição de proficiência da 8ª série (9º ano) deste ano foram, respectivamente, 250 e 50, atribuídos de forma arbitrária. Esta escala foi baseada na estimação conjunta dos parâmetros dos itens de todas as séries das avaliações aplicadas em 1995 e 1997, utilizando a 8ª série de 1997 como grupo de referência e todas as respostas individuais de todos os alunos das três séries nos anos de 1995 e 1997.

Na avaliação aplicada em 1999 a equalização entre anos para a mesma disciplina e a estimação dos parâmetros dos itens da avaliação aplicada neste ano na mesma escala de 1997, foram obtidos conjuntamente com a calibração entre séries, sendo incluída todas as respostas individuais de todos os alunos da amostra de 1997. A 8ª série de 1997 foi novamente utilizada como grupo de referência.

A equalização entre anos e a estimação dos parâmetros da avaliação aplicada pelo SAEB em 2001 na mesma escala de 1997, foram obtidos também conjuntamente com a calibração entre as séries, desta vez acrescentando-se todas as respostas individuais dos alunos das três séries da amostra de 1999 e a 8ª série deste ano foi utilizada como grupo de referência, isto foi possível, pois a avaliação aplicada em 2001 tinha itens comuns com a avaliação aplicada em 1999 em todas as séries avaliadas. Neste ano (2001) foi utilizada a escala (0, 1) para o grupo de referência, como os parâmetros dos itens de todas as séries de 1999 já eram conhecidos, estimados na escala (250, 50), foi feita uma transformação nos parâmetros desses itens para a escala (0, 1). Após a calibração foi feita novamente a transformação para que os parâmetros dos itens (aplicados em 1999) voltassem ao valor original e os parâmetros dos itens novos ficassem na escala comum (250, 50), adotada pelo SAEB.

Para a estimação das proficiências dos alunos, foi pressuposto uma distribuição a priori para as proficiências dos alunos de cada grupo, a distribuição adotada pelo SAEB foi a distribuição normal com média zero e desvio padrão um, para todos os grupos. No processo usado para estimar os parâmetros dos itens obtém-se a distribuição a posteriori de proficiência, condicional às respostas dadas no teste para cada aluno. Assim, a proficiência do aluno foi calculada utilizando-se a média dessa distribuição a posteriori, conhecida como EAP (expected a posteriori). Sendo este, o método para estimação das proficiências dos alunos utilizado atualmente pelo SAEB em todas as suas avaliações.

Em relação à interpretação da escala, sabe-se que é necessário a seleção de alguns níveis da escala, os chamados níveis âncora e para estes são identificados itens âncora, de acordo com alguma metodologia, estes por sua vez são utilizados para a interpretação da escala. Nos anos de 1995 e 1997 foi utilizada a metodologia indicada por Beaton e Allen (1992). Com a utilização dessa metodologia foram observados problemas tais como:

- A necessidade de se escolher níveis âncora muito espaçados a fim de conseguir muitos itens âncora;
- A necessidade de ampliar o conceito de nível âncora para quase âncora (item não satisfaz as três condições) e assim conseguir mais itens facilitando a interpretação da escala;
- A dificuldade dos especialistas de utilizarem o conceito de discriminação entre níveis só usando praticamente itens considerados dominados pelos alunos em torno do nível;

- Alguns descritores cobertos por alguns itens não aparecerem na descrição, pois mesmo sendo bons itens, não eram classificados como itens âncora.

Para contornar esses problemas, no ano de 1999 e também em 2001 foi adotada uma nova metodologia para identificação dos itens âncora. Foram selecionados pontos da escala de 100 a 425, com intervalos de 25 o que incluía o ponto 250 (média adotada em 1997). Feita a estimação do percentual de acerto dos alunos em cada nível para cada item, estes foram posicionados como âncora se atendesse as seguintes condições:

- (i) O número de alunos no nível que respondeu (corretamente ou não) ao item foi maior que 50;
- (ii) O percentual de acerto do item nos níveis anteriores é menor que 65%;
- (iii) O percentual de acerto do item no nível considerado e nos níveis acima é maior que 65%;
- (iv) O ajuste da curva é bom.

Segundo informações fornecidas pelo Serviço de informação ao Cidadão-SIC (informação pessoal)¹ o ajuste da curva é bom quando as proporções de resposta empíricas e esperadas estão em bom ajuste com curva obtida por meio dos parâmetros dos itens. De acordo com Klein (2009), utilizando esta metodologia para identificar os itens âncora, todo item considerado bom será utilizado, pois será âncora em algum nível, a não ser que seja muito difícil. Esta metodologia continua sendo empregada nas avaliações aplicadas pelo SAEB para identificação dos itens âncora.

A escala interpretada no SAEB atualmente tem média 250 e desvio padrão 50 considerando intervalos de comprimento 25 (meio desvio padrão), a fim de se obter uma melhor precisão na descrição pedagógica da escala.

2.2.2 Construção da Escala de Proficiência do ENEM

A TRI começou a ser utilizada no ENEM a partir de 2009, quando foram estabelecidas as matrizes de referência para cada área do conhecimento avaliada. Neste ano foi estabelecida uma escala, a escala “ENEM”, a fim de possibilitar o acompanhamento e comparação dos participantes ao longo dos anos. Para a construção desta escala, foram considerados os participantes concluintes do EM regular de 2009, como grupo de referência e foram definidos a média e o desvio padrão desse grupo como sendo 500 e 100 respectivamente.

No ENEM, de acordo com informações fornecidas através do SIC (informação pessoal)², o modelo da TRI utilizado também é o logístico de três parâmetros. Durante as análises, primeiro são estimados os parâmetros dos itens para, em seguida, estimar-se as proficiências dos participantes. O método utilizado para a estimação dos parâmetros dos itens, assim como no SAEB, é o de Máxima Verossimilhança Marginal. A estimação das proficiências tem como base o método bayesiano EAP. A implementação do modelo também é feita utilizando o software BILOG-MG.

As proficiências na TRI são estimadas em uma escala que não possui mínimo ou máximo pré-estabelecidos, variando de acordo com as características que contém os itens que compõem o exame aplicado a cada edição, isto é, o mínimo e o máximo no ENEM varia em cada edição da prova. Assim sendo, se um aluno acertar todos os itens no exame do ENEM, não significa que este terá uma proficiência igual a 1000. Do mesmo modo

¹Informação fornecida pelo SIC através do protocolo 23480-006786/2018-56 em 29 mar. 2018.

²Informação fornecida pelo SIC através do protocolo 23480.028994/2017-25 em 18 jan. 2018.

que, um participante que erra todas as questões não terá proficiência igual a zero, este recebe o valor mínimo do exame. (BRASIL, 2017e).

2.2.3 Construção das Escalas de Proficiência do SARESP

A TRI vem sendo utilizada nas avaliações do SARESP desde de sua implantação em 1996. Nos primeiros anos de aplicação, a avaliação era realizada ano a ano em cada série/ano de maneira isolada com itens totalmente distintos de um ano para o outro, impedindo o processo de equalização, que tornaria possível a comparação entre anos, proporcionando a verificação de ganho de conhecimento de uma série/ano para a seguinte. Para resolver esse problema foi criada uma prova de ligação, composta por itens que haviam sido submetidos às duas populações (de um ano e do ano seguinte). Com a utilização dessas provas de ligação foi possível a criação de uma escala única para as séries/anos consecutivos, permitindo assim a comparação dos resultados e a criação de escalas de proficiência interpretáveis para cada disciplina avaliada. O trabalho publicado por Valle (2000) traz uma discussão detalhada de como foi feito o processo de equalização das avaliações aplicadas nos anos de 1996, 1997 e 1998 utilizando a prova de ligação.

Atualmente, de acordo com Saresp (2017), a análise dos resultados das avaliações aplicadas pelo SARESP é realizada por meio da análise clássica dos itens e da TRI. A análise clássica, implementada no ItemAN, plataforma de software que automatiza a criação de relatórios de análise de itens com base na teoria clássica de testes, é utilizada para compor o relatório sobre os itens. Uma das medidas utilizadas para análise de discriminação do item, segundo a teoria clássica, é a correlação bisserial, apresentada para cada alternativa do item, é uma medida de associação entre o desempenho no item e o desempenho no teste. Os itens são examinados quando algum distrator (alternativas incorretas) tem coeficiente bisserial maior que a da alternativa correta, quando a alternativa correta tem coeficiente bisserial negativo ou muito próximo de zero ou quando o item tem várias alternativas com coeficientes bisseriais positivos.

A TRI é utilizada para fornecer um estudo mais detalhado sobre os itens e sobre a proficiência dos alunos. Na última prova aplicada pelo SARESP em 2016, o modelo adotado para análise dos parâmetros dos itens de múltipla escolha foi o da Ogiva Normal e para análise das provas com itens de resposta aberta, o modelo adotado foi o de Resposta Gradual. O processo de calibração dos itens e das proficiências utilizando o Modelo Ogiva Normal foi feito pelo software BILOG-MG e para o 3º ano EF, que é constituído de perguntas abertas, o processo de calibração dos itens e das proficiências utilizando o modelo de Resposta Gradual foi utilizado software IRTPro.

O método adotado para a estimação dos parâmetros dos itens foi o de Máxima Verossimilhança Marginal e a estimação das proficiências foi obtida pela média da distribuição a posteriori utilizando a distribuição normal padrão como distribuição da proficiência.

Para definir o posicionamento dos itens na escala foram utilizados os seguintes critérios:

- Itens de resposta aberta - posicionado no nível da escala no qual a probabilidade de resposta na categoria mais alta do nível considerado foi maior ou igual a 0,65;
- Itens de múltipla escolha - Posicionado no nível da escala se a probabilidade de acerto no nível considerado for maior ou igual a 0,65 e a probabilidade de acerto do item no nível dois espaçamentos anterior ao nível considerado foi menor que 0,50.

Os resultados das proficiências obtidas no SARESP para todas as séries/anos avaliados estão na mesma escala do SAEB, com média 250 e desvio padrão 50. Para interpretação das escalas, os níveis são escolhidos como no SAEB com espaçamento de 25 pontos, o equivalente a meia unidade de desvio padrão da escala de proficiência. A Escala de

Matemática, por exemplo, é interpretada nos pontos 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 450 e 475. A descrição de cada um dos pontos é feita com base nos resultados de desempenho dos alunos na prova de Matemática e de acordo com as habilidades detalhadas nas Matrizes de Referência para Avaliação do SARESP. Os pontos da escala, por sua vez, são agrupados em quatro níveis de proficiência, sendo estes, abaixo do básico, básico, adequado e avançado. O agrupamento dos pontos da escala são feitos de acordo com os resultados das proficiências obtidos para cada uma das séries/anos submetidos à avaliação e em cada uma das áreas de conhecimento avaliadas.

A escala está dividida em pontos que vão de 0 (zero) a 500 (quinhentos), porém, é importante ressaltar, novamente, que as escalas de proficiência na TRI não possuem máximo ou mínimo pré-estabelecidos. Estes valores irão variar de acordo com as características dos itens que compõem a avaliação de cada edição.

2.3 Proposta para Construção da Escala a partir de sua Segmentação

Um dos objetivos deste trabalho foi realizar um estudo sobre construções de escalas baseadas na TRI e a partir disto, construir e interpretar uma escala de proficiência em conteúdos matemáticos básicos, adotando-se uma das técnicas estudadas que melhor se adequasse ao contexto abordado. De acordo com o que foi apresentado até o momento, sobre a construção de escalas na TRI e as técnicas adotadas pelos grandes sistemas de avaliação educacional estudados, pôde-se observar que em geral são selecionados níveis âncora na escala adotada, que serão interpretados de acordo com os itens âncora posicionados, seguindo algum critério específico. Desta forma, como já mencionado, não é possível prever o número de itens que serão posicionados como âncora e nem se serão posicionados itens em todos os níveis escolhidos para interpretação. Corre-se o risco de que níveis selecionados não possam ser interpretados.

Buscando melhorar a interpretação da escala em toda a sua extensão, baseado na proposta adotada por Barros (2016), que utilizou a técnica de agrupamentos hierárquicos para construção de uma escala para medida do nível socioeconômico familiar, optou-se por fazer a segmentação da escala classificando os indivíduos em faixas de proficiência, formando grupos e interpretando as faixas dentro de cada grupo. Assim, para atingir este objetivo será aplicada a técnica para segmentação, denominada análise de agrupamentos hierárquicos, a partir do método de Ward. De acordo com Hair et al. (2009), os procedimentos hierárquicos foram os primeiros métodos de agrupamento desenvolvidos, sendo utilizados para classificar indivíduos/elementos em grupos de modo que cada indivíduo é muito semelhante aos indivíduos do grupo que pertence e diferente dos demais grupos.

Com a utilização dessa técnica classifica-se em grupos os níveis selecionados na escala de proficiência e a partir do posicionamento dos itens como âncora nesses níveis, realiza-se a interpretação dentro de cada grupo. Para informações mais detalhadas sobre a análise de agrupamentos hierárquicos e outros métodos de agrupamentos pode-se consultar (HAIR et al., 2009).

Teoria da Resposta ao Item: Conceitos, Definições e Estimação de Parâmetros

O principal foco da TRI, como o próprio nome sugere, é o item e não o teste como um todo, contemplando a análise de itens através das estimativas de seus parâmetros e a análise do instrumento de medida como um todo. Dentro deste contexto, a medida da proficiência de um aluno, por exemplo, não depende dos itens apresentados a ele, e os parâmetros de discriminação e dificuldade do item não dependem do grupo de respondentes. Isto é, na TRI um item mede determinado conhecimento, independentemente de quem o está respondendo bem como, a proficiência de um indivíduo não depende dos itens apresentados a ele (ANJOS; ANDRADE, 2012).

Segundo Anjos e Andrade (2012), a TRI foi desenvolvida com o propósito de superar problemas da Teoria Clássica de Testes, (que utiliza o escore do teste como referência de medida), como a dependência da medida de proficiência em relação ao teste aplicado e a dependência dos parâmetros dos itens em relação ao número de respondentes.

De acordo com Andrade, Tavares e Valle (2000), a TRI consiste em um conjunto de modelos matemáticos, que representam a probabilidade de um indivíduo dar uma resposta certa a um item em função dos parâmetros do item e da proficiência (ou traço latente) do respondente. Essa relação é sempre expressa de tal modo que, quanto maior a proficiência, maior será a probabilidade de resposta correta ao item. Os diversos modelos propostos na literatura dependem fundamentalmente de três fatores:

- (i) da natureza do item, podem ser dicotômicos ou não;
- (ii) do número de populações envolvidas, pode-se ter apenas uma ou mais de uma;
- (iii) da quantidade de traços latentes que está sendo medida, pode-se medir apenas um (modelo unidimensional) ou mais de um (modelo multidimensional).

Dentro do contexto da TRI é importante que se deixe claro o conceito de grupo e de população. De acordo com Andrade, Tavares e Valle (2000) quando se utiliza o termo grupo, este se refere a uma amostra de indivíduos de uma população e quando é feita referência a dois grupos - ou mais - de indivíduos, estes serão dois conjuntos distintos, amostrados de duas - ou mais - populações.

Nas seções que se seguem, serão apresentados, os conceitos, definições e métodos de estimação dos parâmetros e proficiências na TRI. Na Seção 3.1 apresenta-se alguns dos modelos matemáticos presentes na literatura. Na Seção 3.2 o modelo logístico unidimensional de três parâmetros, modelo utilizado nesta pesquisa para a estimação da proficiência em

conteúdos matemáticos básicos de alunos matriculados no primeiro semestre do primeiro ano nos cursos da área de exatas. E por fim, na Seção 3.3 apresenta-se alguns métodos de estimação existentes com ênfase no método bayesiano de estimação dos parâmetros e das proficiências, adotado neste trabalho.

3.1 Modelos Matemáticos

Como citado, a TRI consiste em um conjunto de modelos matemáticos que dependem fundamentalmente: da natureza do item; no número de populações envolvidas e do número de traços latentes que está sendo medido. Assim, encontra-se na literatura, modelos que são utilizados quando um teste é aplicado a um único grupo de respondentes e modelos que são utilizados quando se tem duas ou mais populações. Para os testes aplicados a um único grupo de respondentes, encontra-se modelos utilizados para itens dicotômicos (ou dicotomizados) e modelos utilizados para itens não dicotômicos (ou politômicos). Já para testes aplicados a duas ou mais populações, os modelos utilizados, sugeridos da literatura, são extensões dos modelos aplicados a um único grupo de respondentes.

Os modelos de resposta ao item mais utilizados para itens dicotômicos, são os chamados modelos logísticos que, segundo Andrade, Tavares e Valle (2000), podem ser aplicados tanto para análise de itens de múltipla escolha dicotomizados, ou seja, corrigidos como certo ou errado, quanto para a análise de itens abertos, isto é, de resposta livre, corrigidos de forma dicotomizada. Existem essencialmente três modelos logísticos que se diferem pelo número de parâmetros. São conhecidos como modelo logístico de um parâmetro (ML1) onde é levado em consideração somente a dificuldade do item, de dois parâmetros (ML2) onde são considerados os parâmetros de dificuldade e discriminação e o de três parâmetros (ML3) que considera os parâmetros de dificuldade, discriminação e a probabilidade de acerto ao acaso (chute).

O modelo logístico unidimensional de três parâmetros (ML3) será apresentado com mais detalhes visto que, este foi o modelo utilizado no presente trabalho para avaliar respostas dicotomizadas (corrigidas como certo ou errado), com o intuito de medir apenas um traço latente (ou proficiência), por esta razão, chamado de modelo unidimensional. A saber, mediu-se a proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo e similares dos alunos matriculados no primeiro semestre do primeiro ano nos cursos da área de exatas, utilizando como instrumento de medida uma prova composta por itens de múltipla escolha.

Quando se tem itens classificados como não dicotômicos, são encontrados na literatura modelos para análise de itens abertos (os de resposta livre) e modelos para análise de itens de múltipla escolha que são avaliados de forma graduada, itens deste tipo são construídos ou corrigidos de forma que se tenha uma ou mais categorias intermediárias ordenadas entre as categorias certo ou errado. Este tipo de item não leva-se em conta somente se o indivíduo respondeu de forma correta ou não a alternativa, considera-se também qual foi a resposta dada por ele. Os modelos de Resposta Nominal, de Resposta Gradual, de Escala Gradual, de Crédito Parcial e o modelo de Crédito Parcial Generalizado são exemplos de modelos utilizados para itens não dicotômicos.

O Modelo de Resposta Nominal pode ser aplicado a todas as categorias de respostas escolhidas em um teste com itens de múltipla escolha. Este modelo tem o objetivo de potencializar a precisão do traço latente estimado usando toda a informação contida nas respostas dos indivíduos e não somente se o item foi respondido corretamente ou não. O modelo de Resposta Gradual admite que as categorias de resposta de um item possam ser ordenadas entre si, tendo este o mesmo propósito do modelo de Resposta Nominal. O modelo de Escala Gradual corresponde a um caso particular do Modelo de Resposta

Gradual, que também é aplicado a itens com categoria de respostas ordenadas. Além disso, supõe-se que os escores das categorias são igualmente espaçados. O modelo de Crédito Parcial também pode ser utilizado para análise de respostas adquiridas de categorias ordenadas e tem o mesmo propósito dos modelos anteriores. O modelo de Crédito Parcial Generalizado, também indicado para análise de itens não dicotômicos, como citado, é baseado no modelo de Crédito Parcial. Informações mais detalhadas sobre esses modelos podem ser obtidas em (ANDRADE; TAVARES; VALLE, 2000).

Para testes aplicados a duas ou mais populações são encontradas na literatura extensões dos modelos aplicados a um único grupo de respondentes. De acordo com Andrade, Tavares e Valle (2000) generalizações dos modelos logísticos unidimensionais de 1, 2 e 3 parâmetros são utilizados para um teste aplicado a dois ou mais grupos de respondentes. O que acontece em geral é que indivíduos pertencentes a diferentes populações não são submetidos aos mesmos testes. Mas, para que seja possível efetuar comparações entre as populações é necessário haver alguns itens comuns entre eles.

A utilização de itens comuns em testes diferentes aplicados a populações diferentes, possibilita que ao final do processo de estimação, todos os parâmetros (dos itens e também das proficiências) estejam em uma mesma escala, viabilizando comparações entre as populações e a construção de escalas de proficiência (ou escalas de conhecimento) que possam ser interpretadas, o que é de grande importância para área educacional. O processo que possibilita a inclusão de todos os parâmetros, tanto dos itens quanto das proficiências em uma mesma métrica é conhecido como processo de equalização. O processo de equalização pode ser feito de duas maneiras diferentes. O caso mais recorrente acontece em populações distintas que respondem provas parcialmente diferentes. Informações mais detalhadas sobre este processo de equalização e outros tipos diferentes de equalizações podem ser vistos em (ANDRADE; TAVARES; VALLE, 2000).

3.2 O Modelo Logístico Unidimensional de três Parâmetros (ML3)

Como mencionado na Seção 3.1, os modelos mais utilizados para itens dicotômicos são os modelos logísticos. O modelo logístico de dois parâmetros foi modificado por Birnbaum (1968) para incluir o parâmetro c que representa a probabilidade de acerto ao acaso. O modelo resultante tornou-se conhecido como modelo logístico de três parâmetros (ML3) que é considerado o mais completo, sendo possível obter os modelos de 1 e 2 parâmetros facilmente a partir deste. O modelo logístico unidimensional de três parâmetros (ML3) conforme citado, foi o modelo utilizado neste trabalho, pois os itens que fazem parte da prova aplicada aos alunos matriculados no primeiro semestre do primeiro ano dos cursos da área de exatas, serão analisados de maneira dicotômica (corrigidos como certo ou errado) e utilizados para avaliar apenas um traço latente de interesse nesta pesquisa que é a proficiência em conteúdos matemáticos básicos.

De acordo com Andrade, Tavares e Valle (2000) o modelo (ML3) é dado por

$$P(U_{ij} = 1 | \theta_j) = c_i + (1 - c_i) \frac{1}{1 + \exp^{-Da_i(\theta_j - b_i)}} \quad (3.1)$$

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

- I representada o número de item que compõem o teste;
- n representa o número de respondentes;

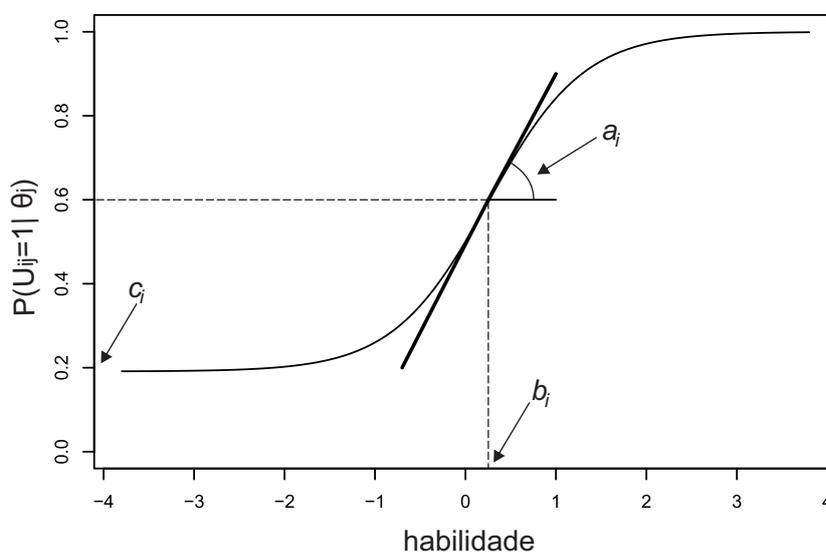
- U_{ij} corresponde a uma variável dicotômica que assume os valores, 1 quando o respondente j responde corretamente o item i , ou 0 caso contrário;
- θ_j corresponde ao traço latente (ou proficiência) do j -ésimo indivíduo;
- $P(U_{ij} = 1 | \theta_j)$ corresponde a probabilidade de um indivíduo j com traço latente θ_j responder corretamente o item i , chamada de Função de Resposta do Item - FRI;
- b_i corresponde ao parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala do traço latente;
- a_i corresponde ao parâmetro de discriminação (ou inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item - CCI no ponto b_i ;
- c_i corresponde ao parâmetro do item que representa a probabilidade de indivíduos com baixa proficiência responderem corretamente a um item i (muitas vezes mencionada como a probabilidade de acerto casual);
- D corresponde a um fator de escala, constante e igual a 1. Quando se desejar que a função logística forneça resultados semelhantes aos da função ogiva normal, utiliza-se para o fator D o valor 1,7.

3.2.1 Curva Característica do Item - CCI: Interpretação e Representação Gráfica

A curva característica do item possui formato em “S” e descreve a relação entre a probabilidade de resposta correta a um item e a escala de proficiência. De acordo com Andrade, Tavares e Valle (2000), a escala de proficiência é uma escala arbitrária onde o que importa são as relações de ordem existentes entre seus pontos e não necessariamente a sua magnitude.

A relação entre a probabilidade de um indivíduo j com proficiência θ_j responder corretamente a um item i , dada por $P(U_{ij} = 1 | \theta_j)$, e a escala de proficiência, bem como os parâmetros do modelo e sua forma em “S”, podem ser visualizadas na Figura 3.1. Pode-se observar que para valores mais baixos de proficiência, a probabilidade de resposta correta se encontra nos níveis mais baixos. Para indivíduos com maiores proficiências as probabilidades de resposta correta é maior. Observe que essa relação é não linear. Baker (2001) enfatiza que a Curva Característica do Item é alicerce de construção básico da TRI e que dela dependem todas as outras construções, razão pela qual ele afirma que deve ser dada a esta curva considerável atenção dentro da teoria.

Figura 3.1: Exemplo de uma curva característica do item - CCI.



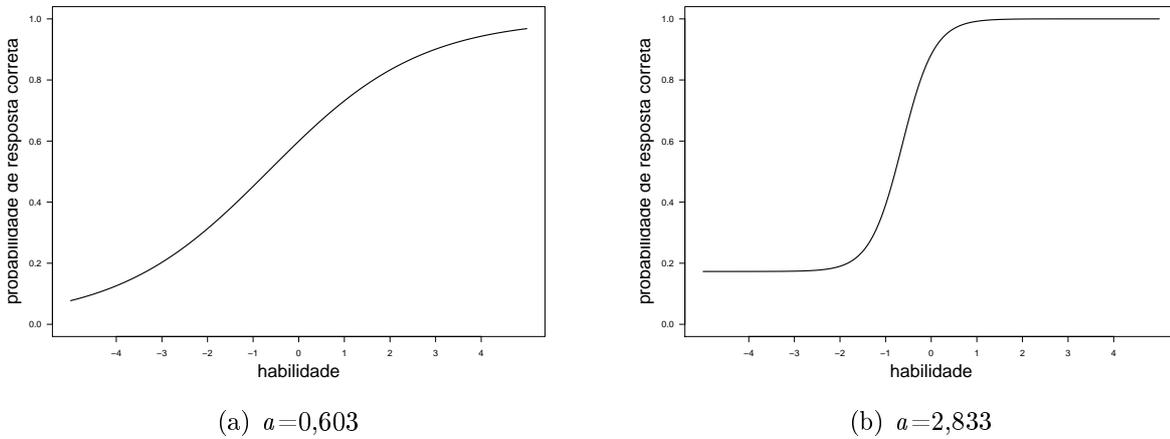
Fonte: Elaborado pela autora (adaptado de (ANDRADE; TAVARES; VALLE, 2000)).

O parâmetro a é proporcional à derivada da tangente da curva no ponto de inflexão. Portanto, não são esperados itens com valores de a negativo para este modelo, pois isto indicaria que a probabilidade de responder corretamente o item diminui com o aumento da proficiência, o que seria contraditório, uma vez que o modelo proposto baseia-se no fato de que indivíduos com maior proficiência tem maior probabilidade de acertar o item.

Se o parâmetro a possui um valor baixo, obtêm-se curvas características com uma inclinação menor, o que indica que o item tem pouco poder de discriminação. Quando isso acontece, indivíduos com proficiência baixa e indivíduos com proficiência alta terão aproximadamente a mesma probabilidade de acertar o item. Se o parâmetro a possui um valor alto, obtêm-se curvas características “íngremes”, que acabam dividindo os indivíduos em dois grupos: os que possuem proficiência abaixo da localização do item e os que possuem proficiência acima dessa localização.

Isso pode ser observado na Figura 3.2. O exemplo (a) traz a CCI de um item com parâmetro a de valor mais baixo, menos discriminativo. Pode-se observar que um indivíduo com proficiência 0 por exemplo, tem probabilidade de responder corretamente o item bem próxima a probabilidade do indivíduo com proficiência 2. O exemplo (b) mostra a curva de um item com valor de a alto, bem discriminativo e apresenta uma CCI “íngreme”. Os dois exemplos (a) e (b) tem parâmetro de dificuldade $b = -0.638$, posição onde o item está localizado na escala. Pode-se observar que a maior parte dos indivíduos com proficiência abaixo deste valor em (b), tem a mesma probabilidade de responder corretamente ao item sendo esta, a probabilidade de acerto ao acaso para o item.

Figura 3.2: Exemplos de curva característica do item com valores diferentes para o parâmetro de discriminação.

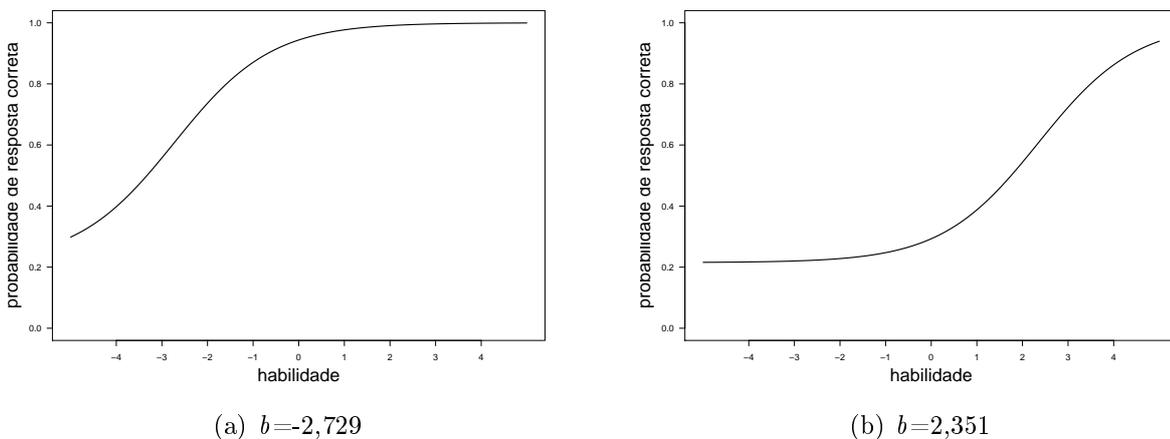


Fonte: Elaborado pela autora.

O parâmetro b é medido na mesma escala da proficiência θ , representando a proficiência necessária para uma probabilidade de acerto dada por $(1 + c)/2$ (ANDRADE; TAVARES; VALLE, 2000). Deste modo, quanto maior o valor de b mais difícil será o item e quanto menor o valor do parâmetro, mais fácil o item será considerado.

A Figura 3.3 mostra dois exemplos de CCI's, onde o valor do parâmetro a e c são os mesmos para os dois exemplos, 0,938 e 0,215 respectivamente. No exemplo (a), a curva está representando um item fácil, podemos observar que a probabilidade de resposta correta é alta já para níveis de proficiência baixos. No exemplo (b), a curva representa um item difícil, podemos observar que só apresenta probabilidade alta de resposta correta, indivíduos com níveis muito altos de proficiência.

Figura 3.3: Exemplos de curva característica do item com valores diferentes para o parâmetro de dificuldade.



Fonte: Elaborado pela autora.

O parâmetro c representa a probabilidade de um indivíduo com baixa proficiência responder corretamente a um item, essa probabilidade pode ser chamada de probabilidade de acerto ao acaso. Segundo Andrade, Tavares e Valle (2000) quando não é permitido “chutar”, o parâmetro c é igual a 0 e o parâmetro b representa o ponto na escala de proficiência onde a probabilidade de acertar o item é de 0.5.

3.2.2 Função de Informação do Item

A função de informação do item é uma medida bastante utilizada em conjunto com a CCI. Permite analisar quanto um item contém de informação para a medida de proficiência. Conforme apresentado por Andrade, Tavares e Valle (2000) a função de informação de um item é dada por

$$I_i(\theta) = \frac{\left[\frac{\partial}{\partial \theta} P_i(\theta) \right]^2}{P_i(\theta)Q_i(\theta)}, \quad (3.2)$$

onde: $I_i(\theta)$ é a informação fornecida pelo item i no nível de proficiência θ ; $P_i(\theta) = P(U_{ij} = 1 | \theta)$; $Q_i(\theta) = 1 - P_i(\theta)$ e $\frac{\partial}{\partial \theta}$ é a derivada da função em relação a θ .

Para o modelo logístico de três parâmetros (ML3), a equação pode ser escrita como:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i} \right]^2. \quad (3.3)$$

de fato,

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

Segundo Baker (2001) a função de informação do item de um modelo logístico de três parâmetros (ML3), com os mesmos valores para os parâmetros a e b de um modelo logístico de dois parâmetros (ML2), oferece uma quantidade menor de informação, o que é aceitável, pois obter o item correto por adivinhação não deve aumentar a precisão com que um nível de proficiência é estimado. O máximo da função de informação também

não ocorre a um nível de proficiência equivalente ao parâmetro de dificuldade b . Sendo assim, a equação (3.3) mostra a importância que têm os três parâmetros sobre o total de informação do item. A informação será maior quando:

- (i) b_i se aproxima de θ ;
- (ii) maior for o a_i ;
- (iii) mais c_i se aproximar de 0.

3.2.3 Função de Informação do Teste

De acordo com Baker (2001) uma vez que se utiliza um teste para estimar a proficiência de um indivíduo, a quantidade de informação produzida pelo teste para qualquer nível de proficiência também pode ser obtida. Um teste é composto por um conjunto de itens, assim a informação fornecida pelo teste em um determinado nível de proficiência é simplesmente a soma das informações dos itens nesse nível. Desta forma, a informação fornecida pelo teste é dada pela soma das informações fornecidas pelos itens que o compõe.

Segundo Andrade, Tavares e Valle (2000) a função de informação do teste é dada por

$$I(\theta) = \sum_{i=1}^I I_i(\theta) \quad (3.4)$$

onde,

- (i) $I(\theta)$ representa a quantidade de informação fornecida pelo teste no nível de proficiência θ ;
- (ii) $I_i(\theta)$ representa a quantidade de informação fornecida pelo i -ésimo item no nível de proficiência θ ;
- (iii) I representa a quantidade de itens que compõe o teste.

A função de informação do teste tem a característica de expressar o quão bem o teste está estimando as proficiências em toda a gama de pontuação de proficiências. Pode-se notar que a medida de informação em (3.4) depende do valor de θ . Assim, a amplitude do intervalo de confiança para θ também dependerá do seu valor.

3.2.4 Suposições do Modelo

As suposições consideradas fundamentais em modelos da TRI são a unidimensionalidade e a independência local embora, segundo Nojosa (2002) a independência local não é mais vista como uma suposição e sim como uma consequência da unidimensionalidade.

Para o modelo proposto a suposição de unidimensionalidade do teste exige que apenas um único traço latente (ou proficiência) esteja sendo medido pelo conjunto de itens. Essa suposição não pode ser estritamente cumprida, uma vez que para a execução de qualquer tarefa uma pessoa põe em prática mais de um traço latente. Fatores, tais como o nível de motivação, a ansiedade, a capacidade de trabalhar rapidamente, a tendência de adivinhar quando em dúvida sobre respostas são exemplos de habilidades que podem ser utilizadas conjuntamente com a proficiência que se quer medir quando um indivíduo é submetido a um teste. Para que o postulado de unidimensionalidade seja satisfeito é necessário admitir que existe uma proficiência dominante (ou fator dominante) responsável pelo

conjunto de itens. Esta proficiência dominante é o que se pretende medir no teste. Os desvios que as habilidades secundárias produzem na interpretação dos escores de um teste são considerados suficientemente pequenos (ANDRADE; TAVARES; VALLE, 2000), (PASQUALI, 2003) e (HAMBLETON; SWAMINATHAN; ROGERS, 1991).

A suposição de independência local assume que para uma determinada proficiência, as respostas aos diferentes itens que compõe um teste são independentes. De acordo com Andrade, Tavares e Valle (2000) a unidimensionalidade implica na independência local, desta forma é necessário verificar apenas uma das suposições, sendo a suposição de independência local fundamental para o processo de estimação dos parâmetros.

Uma outra suposição que também deve ser considerada é a independência das respostas dos indivíduos, ou seja, as respostas dos indivíduos diferentes são consideradas independentes.

3.3 Estimação dos Parâmetros

A estimação dos parâmetros em geral é realizada a partir dos processos de inferência clássico ou bayesiano. Nas seções que se seguem é apresentada uma breve descrição de métodos de estimação presentes na literatura utilizados na abordagem clássica e a estimação sob enfoque bayesiano, adotado nesta pesquisa.

3.3.1 Uma Breve Descrição dos Métodos Propostos na Literatura

Existem dois tipos de parâmetros que caracterizam os modelos TRI, os parâmetros dos itens e das proficiências. Os parâmetros dos itens estão relacionados às questões e o parâmetro da proficiência está relacionado aos indivíduos. Desta forma, nos modelos TRI temos um problema de estimação que envolve os dois tipos de parâmetros, os dos itens e os das proficiências dos indivíduos, podendo ocorrer três situações diferentes: A situação em que os parâmetros dos itens já estão estimados e deseja-se estimar somente as proficiências dos indivíduos; o caso onde se conhece as proficiências dos indivíduos e deseja-se estimar os parâmetros dos itens e a situação em que tanto os parâmetros dos itens quanto das proficiências são desconhecidos. Na TRI constantemente nos referimos ao processo de estimação dos parâmetros dos itens como calibração dos itens (ANDRADE; TAVARES; VALLE, 2000); (AZEVEDO, 2003).

Andrade, Tavares e Valle (2000) enfatizam que a etapa da estimação dos parâmetros é uma das mais importantes da TRI, dentro do contexto da abordagem clássica, normalmente a estimação é feita pelo Método da Máxima verossimilhança (MV) sendo necessária a utilização de algum processo iterativo, como o algoritmo de *Newton-Raphson* ou *Scoring de Fisher*. Para o caso mais comum, onde há a necessidade de estimar tanto os parâmetros dos itens quanto das proficiências, segundo Andrade, Tavares e Valle (2000) tem-se duas abordagens usuais, a estimação conjunta dos parâmetros dos itens e das proficiências e a estimação em duas etapas, primeiro a estimação dos parâmetros dos itens e depois das proficiências. No processo de estimação conjunta, chamado de Máxima Verossimilhança conjunta (MVC) ocorre que o número de parâmetros a serem estimados simultaneamente pode ser extremamente grande exigindo um grande esforço computacional e pode acarretar problemas de convergência.

Na tentativa de contornar esse problema, um processo de *vai e volta* ("*back-and-forth*") foi proposto por Birnbaum (1968). O processo é iniciado com estimativas grosseiras para as proficiências e então é feita a estimação dos parâmetros dos itens, considerando as proficiências conhecidas. Obtendo as estimativas dos parâmetros dos itens, estima-se as proficiências considerando conhecidos os parâmetros dos itens. Esses passos são

repetidos até que algum critério de convergência seja atingido. Segundo Andrade, Tavares e Valle (2000) a vantagem desse método é que levando em conta a independência local, os itens podem ser estimados individualmente e levando em conta a independência entre as respostas oriundas de indivíduos diferentes, as proficiências também podem ser estimadas individualmente, o que diminuirá consideravelmente o esforço computacional.

Sabe-se que, quando os parâmetros dos itens são conhecidos, os Estimadores de Máxima Verossimilhança (EMV) das proficiências convergem para os seus verdadeiros valores quando o número de itens cresce. Da mesma forma, quando as proficiências são conhecidas, os EMV dos parâmetros dos itens convergem para seus verdadeiros valores quando o número de indivíduos cresce. De acordo com Andrade, Tavares e Valle (2000) na estimação conjunta, os parâmetros das proficiências são chamados de parâmetros incidentais, pois se a amostra cresce o número de parâmetros θ_j também cresce e os parâmetros dos itens são chamados de parâmetros estruturais, pois se o número de indivíduos cresce o número de parâmetros dos itens não se altera. Deste modo, um sério problema deste procedimento é que EMV dos parâmetros dos itens (ou proficiências) na presença de um número muito grande de indivíduos (ou itens) pode ser assintoticamente viciado. No entanto, o EMV dos parâmetros dos itens e das proficiências serão não-viciados, quando o número de itens e o número de indivíduos crescem.

Para contornar este problema de inconsistência dos estimadores obtidos em uma etapa, Bock e Lieberman (1970) propuseram um método de estimação conhecido como Método da Máxima Verossimilhança Marginal (MVM), obtido em duas etapas. De acordo com Andrade, Tavares e Valle (2000) a ideia consiste em considerar uma distribuição chamada de latente, associada às proficiências de uma população II em estudo, composta por n indivíduos com proficiências θ_j , $j = 1, \dots, n$ e utilizar uma verossimilhança marginal, integrando a função de verossimilhança com relação ao parâmetro θ , para estimar os parâmetros dos itens. Estimado os parâmetros dos itens, as proficiências são estimadas individualmente por máxima verossimilhança, ou pela moda ou pela média da distribuição condicional de θ_j .

O método possui a vantagem de envolver, na primeira etapa, apenas a estimação dos parâmetros dos itens e gera estimativas de MV com propriedades assintóticas satisfatórias. Mas, ainda requer a estimação conjunta dos parâmetros de todos os itens através de métodos numéricos, exigindo grande esforço computacional.

Bock e Aitkin (1981) tentando contornar esta dificuldade, propuseram uma modificação do MVM proposto por Bock e Lieberman (1970), acrescentando a suposição de independência entre os itens, de maneira que os parâmetros dos itens possam ser estimados individualmente, reduzindo o esforço computacional e preservando as propriedades assintóticas dos estimadores. Além disso, propuseram também que as estimativas de MV fossem obtidas com aplicação do algoritmo EM. Proposto por Dempster, Laird e Rubin (1977) consiste em um processo iterativo para determinação de EMV's de parâmetros de modelos de probabilidade com a presença de variáveis aleatórias não observadas, onde cada iteração do processo é feita em dois passos. Um passo de Esperança e um passo de Maximização. Informações mais detalhadas sobre o algoritmo EM podem ser consultadas (DEMPSTER; LAIRD; RUBIN, 1977).

Apesar do Método de Máxima Verossimilhança Marginal ter proporcionado um grande avanço nos ajustes de modelos na TRI, os EMV's apresentam problemas na estimação dos parâmetros dos itens que são respondidos corretamente ou incorretamente por todos os indivíduos e na estimação das proficiências de indivíduos que respondem corretamente ou incorretamente todos os itens. Também existe a possibilidade de que as estimativas dos parâmetros dos itens caiam fora do intervalo esperado para cada um, como: valores do parâmetro a_i negativos, ou ainda valores do parâmetro c_i fora do intervalo $[0, 1]$. Os

autores a seguir podem ser consultados para melhor entendimento dos métodos de estimação citados acima, sendo esses os mais utilizados na abordagem clássica (ANDRADE; TAVARES; VALLE, 2000); (AZEVEDO, 2003); (BAKER; KIM, 2004).

A abordagem bayesiana vem sendo utilizada por pesquisadores em TRI com a intenção de aprimorar a precisão das estimativas dos parâmetros dos itens e proficiências, oferecendo uma solução para os problemas apresentados na estimação por máxima verossimilhança. De acordo com Andrade, Tavares e Valle (2000), um dos métodos utilizados é o Método de Estimação Bayesiana Marginal proposta por Mislevy (1986) baseando-se na proposta de Bock e Aitkin (1981). Marques (2008) se refere a este método como uma estimação parcialmente bayesiana, por se tratar de uma extensão da proposta de Bock e Aitkin (1981). Para maiores detalhes sobre o método pode se consultar (ANDRADE; TAVARES; VALLE, 2000).

A estimação dos parâmetros dos itens no software BILOG-MG é feita por Máxima Verossimilhança Marginal, com a especificação de distribuições a priori para os parâmetros dos itens e para as proficiências dos respondentes. Como mencionado, este é o software utilizado pelos três grandes sistemas de avaliação educacional citados neste trabalho. Para mais informações sobre o software BILOG-MG pode-se consultar (DUTOIT, 2003).

3.3.2 Estimação Bayesiana

Para o desenvolvimento dos métodos de estimação bayesianos se faz necessário considerar os seguintes procedimentos: especificar a função de verossimilhança; estabelecer distribuições a priori para os parâmetros de interesse, obter uma nova função, com a utilização do teorema de Bayes, a partir da combinação das informações fornecidas pelas distribuições a priori e a função de verossimilhança, que é chamada de distribuição a posteriori. A forma usual do Teorema de Bayes, que fornece a função densidade de probabilidade a posteriori, é dada por

$$p(\boldsymbol{\lambda}|\mathbf{u}) \propto L(\boldsymbol{\lambda}; \mathbf{u})p(\boldsymbol{\lambda}), \quad (3.5)$$

onde $\boldsymbol{\lambda}$ representa o vetor de parâmetros do modelo ML3, $\boldsymbol{\lambda} = (\theta, a, b, c)$, \propto o símbolo denotando proporcional a, $L(\boldsymbol{\lambda}; \mathbf{u})$ representa a função de verossimilhança e $p(\boldsymbol{\lambda})$ representa a distribuição de probabilidade a priori conjunta.

O teorema de Bayes fornece a regra de atualização de probabilidades sobre $\boldsymbol{\lambda}$, partindo-se de $p(\boldsymbol{\lambda})$ e chegando-se em $p(\boldsymbol{\lambda}|\mathbf{u})$. Por este motivo são chamadas, respectivamente, de distribuição de probabilidade a priori e distribuição de probabilidade a posteriori, como pode ser visto em Migon, Gamerman e Louzada (2014).

Devido à complexidade dos modelos TRI, para o procedimento de inferência sobre os parâmetros dos itens e proficiências, sob o paradigma bayesiano, se faz necessário o uso de métodos numéricos. Os métodos numéricos mais usuais são os Métodos de Monte Carlo via Cadeias de Markov (MCMC). O algoritmo de Metropolis-Hastings e o amostrador de Gibbs são os algoritmos da classe MCMC mais utilizados. Neste trabalho foi utilizado o processo inferência bayesiana para estimação dos parâmetros dos itens que compõem a prova e para a estimação das proficiências dos alunos participantes. O algoritmo utilizado foi o amostrador de Gibbs que consiste em um esquema de gerações sucessivas, a partir de distribuições conhecidas por distribuições condicionais completas, permitindo a obtenção de amostras da distribuição a posteriori (GAMERMAN; LOPES, 2006).

3.3.3 Descrição da Função de Verossimilhança

A fim de se descrever a função de verossimilhança, considere θ_j a proficiência e U_{ji} a variável aleatória que representa a resposta do indivíduo j ao item i , com

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

Pode-se observar que a variável aleatória U_{ji} segue uma distribuição de Bernoulli, $U_{ji} \sim Ber(p_{ji})$, sendo p_{ji} a probabilidade do indivíduo j responder corretamente ao item i .

Considere ainda, $\mathbf{U}_j = (U_{j1}, U_{j2}, \dots, U_{jI})$ o vetor aleatório de respostas do indivíduo j , $\mathbf{U}_{..} = (\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n)$ o conjunto de todas as respostas, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$ o vetor de proficiências dos n indivíduos, $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_I)$ o conjunto de parâmetros dos itens, com $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)$ onde $i = 1, \dots, I$. Por conveniência, as observações serão representadas por u_{ji} , \mathbf{u}_j e $\mathbf{u}_{..}$.

Devido às suposições feitas para o modelo ML3, de independência local e de independência entre as repostas de diferentes indivíduos, a função de verossimilhança $L(\boldsymbol{\theta}, \boldsymbol{\zeta}; \mathbf{u}_{..}) = P(\mathbf{U}_{..} = \mathbf{u}_{..} \mid \boldsymbol{\theta}, \boldsymbol{\zeta})$ pode ser escrita como

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\zeta}; \mathbf{u}_{..}) &= \prod_{j=1}^n P(\mathbf{U}_j = \mathbf{u}_j \mid \theta_j, \boldsymbol{\zeta}) \\ &= \prod_{j=1}^n \prod_{i=1}^I P(U_{ji} = u_{ji} \mid \theta_j, \boldsymbol{\zeta}_i). \end{aligned} \tag{3.6}$$

Usando o fato de que a distribuição U_{ji} só depende de $\boldsymbol{\zeta}$ através de $\boldsymbol{\zeta}_i$ e fazendo $P_{ji} = P(U_{ji} = 1 \mid \theta_j, \boldsymbol{\zeta}_i)$ e $Q_{ji} = 1 - P_{ji}$, temos que

$$\begin{aligned} P(U_{ji} = u_{ji} \mid \theta_j, \boldsymbol{\zeta}_i) &= P(U_{ji} = 1 \mid \theta_j, \boldsymbol{\zeta}_i)^{u_{ji}} P(U_{ji} = 0 \mid \theta_j, \boldsymbol{\zeta}_i)^{1-u_{ji}} \\ &= P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}. \end{aligned} \tag{3.7}$$

e, portanto,

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\zeta}; \mathbf{u}_{..}) &= \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} \\ &= \prod_{j=1}^n \prod_{i=1}^I \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\ &\quad \times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}}. \end{aligned} \tag{3.8}$$

3.3.4 Descrição da Distribuição a Priori

A utilização de informação a priori em Inferência Bayesiana, requer a especificação de uma distribuição para uma ou mais quantidades de interesse desconhecidas. Essa distribuição deve representar o conhecimento que se tem sobre essa quantidade de interesse antes de se realizar o experimento. Esse conhecimento prévio é representado matematicamente pela distribuição de probabilidade que recebe o nome de distribuição a priori,

estabelecendo quais possíveis valores para quantidade de interesse são mais prováveis de acordo com informação que se tem disponível, antes de se conhecer as observações. Essas informações podem ser por exemplo, dados históricos sobre o problema ou até mesmo dados sobre problemas similares.

Para a aplicação da inferência bayesiana no processo de estimação dos parâmetros dos itens e das proficiências nos modelos TRI, além de atribuir uma distribuição de probabilidade a priori para variável latente θ_j , também é necessário atribuir distribuições de probabilidade a priori para os parâmetros dos itens. Serão descritas a partir de agora, as distribuições a priori mais indicadas na literatura e que foram utilizadas neste trabalho para o modelo ML3. As sugestões de utilização das mesmas, podem ser consultadas por exemplo, em (ANDRADE; TAVARES; VALLE, 2000);(MARQUES, 2008);(BAKER; KIM, 2004).

Para a distribuição a priori da variável latente θ_j , $j = 1, \dots, n$, sendo n o número de indivíduos, é assumida a distribuição a priori Normal com hiperparâmetros $\boldsymbol{\eta} = (\mu_\theta, \sigma_\theta^2)$. Sendo sua densidade dada por

$$g(\theta_j | \boldsymbol{\eta}) = \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp \left\{ -\frac{1}{2} \left(\frac{\theta_j - \mu_\theta}{\sigma_\theta} \right)^2 \right\}, \quad j = 1, \dots, n. \quad (3.9)$$

Para o parâmetro a_i são sugeridas a adoção das distribuições *Log-normal* ou a *Chi-quadrado*. Uma justificativa para a adoção dessas distribuições comentada por Andrade, Tavares e Valle (2000) é o fato de que, na prática, os a_i são positivos, sugerindo que a distribuição de a_i pode ser modelada por uma distribuição unimodal e com assimetria positiva, como acontece nestas distribuições. Outra distribuição sugerida na literatura é a normal truncada no zero conforme Curtis et al. (2010) e Fox (2010). Após alguns testes esta última sugestão foi adotada. Portanto, optou-se por atribuir para cada parâmetro a_i uma distribuição Normal truncada no zero com hiperparâmetros $\boldsymbol{\tau} = (\mu_a, \sigma_a^2)$, com densidade dada por

$$f(a_i | \boldsymbol{\tau}) = \frac{1}{\sqrt{2\pi}\sigma_a} \exp \left\{ -\frac{1}{2} \left(\frac{a_i - \mu_a}{\sigma_a} \right)^2 \right\} I_{(0,\infty)}, \quad i = 1, \dots, I. \quad (3.10)$$

Para o parâmetro b_i , em geral, é atribuída uma distribuição a priori Normal, devido ao fato dos parâmetros de dificuldade estarem na mesma escala da proficiência. Atribuiu-se então, para cada parâmetro b_i uma distribuição a priori Normal com hiperparâmetros $\boldsymbol{\gamma} = (\mu_b, \sigma_b^2)$, com densidade dada por

$$f(b_i | \boldsymbol{\gamma}) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp \left\{ -\frac{1}{2} \left(\frac{b_i - \mu_b}{\sigma_b} \right)^2 \right\}, \quad i = 1, \dots, I. \quad (3.11)$$

Para o parâmetro c_i , geralmente, é atribuída uma distribuição a priori Beta, pois c_i só pode pertencer ao intervalo $[0, 1]$. Assumiu-se, então, que cada parâmetro c_i segue uma distribuição a priori Beta com hiperparâmetros $\boldsymbol{\delta} = (\alpha, \beta)$, com densidade dada por

$$f(c_i | \boldsymbol{\delta}) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} c_i^{\alpha-1} (1 - c_i)^{\beta-1}, \quad i = 1, \dots, I. \quad (3.12)$$

A média (ou valor esperado) da distribuição Beta pode ser interpretada como a probabilidade de um indivíduo com baixa proficiência responder corretamente a um dado item. O valor esperado da distribuição Beta apresentada é dado por

$$E_i = E[c_i | \alpha, \beta] = \frac{\alpha}{\alpha + \beta}. \quad (3.13)$$

para $i = 1, \dots, I$, onde I representa o número de itens. Desta forma, pode-se definir valores para os parâmetros α e β de modo que E_i tenha um valor satisfatório (valor esperado de probabilidade ao acaso em cada item). A variância da distribuição Beta é dada por

$$Var[c_i | \alpha, \beta] = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}.$$

Como em Marques (2008) pode-se denotar a distribuição a priori conjunta dos parâmetros dos itens por

$$f(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \prod_{i=1}^I h(a_i, b_i, c_i).$$

Supondo independência a priori entre os parâmetros dos itens, temos

$$h(a_i, b_i, c_i) = h_1(a_i)h_2(b_i)h_3(c_i).$$

Assim,

$$h(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{c}) = \prod_{j=1}^n g(\theta_j) \prod_{i=1}^I h_1(a_i)h_2(b_i)h_3(c_i).$$

onde, $g(\theta_j)$ representa a distribuição a priori para a variável latente θ_j .

3.3.5 Descrição da Distribuição a Posteriori

Sabe-se que a distribuição a posteriori é obtida pelo teorema de Bayes, a partir da combinação entre as informações fornecidas pelas distribuições a priori e pela função de verossimilhança (3.8). A fim de obtermos a distribuição a posteriori para o modelo ML3, considerou-se, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$, $\mathbf{a} = (a_1, \dots, a_I)$, $\mathbf{b} = (b_1, \dots, b_I)$, $\mathbf{c} = (c_1, \dots, c_I)$ e $\mathbf{u} = \mathbf{u}_{..}$. Assim,

$$\begin{aligned} p(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{c} | \mathbf{u}) &\propto \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} \prod_{j=1}^n g(\theta_j) \prod_{i=1}^I [h_1(a_i)h_2(b_i)h_3(c_i)] \\ &\propto \prod_{j=1}^n \prod_{i=1}^I \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\ &\quad \times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}} \\ &\quad \times \prod_{j=1}^n \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp \left\{ -\frac{1}{2} \left(\frac{\theta_j - \mu_\theta}{\sigma_\theta} \right)^2 \right\} \\ &\quad \times \prod_{i=1}^I \left[\frac{1}{\sqrt{2\pi}\sigma_a} \exp \left\{ -\frac{1}{2} \left(\frac{a_i - \mu_a}{\sigma_a} \right)^2 \right\} I_{(0,\infty)} \right] \\ &\quad \times \frac{1}{\sqrt{2\pi}\sigma_b} \exp \left\{ -\frac{1}{2} \left(\frac{b_i - \mu_b}{\sigma_b} \right)^2 \right\} \\ &\quad \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} c_i^{\alpha-1} (1 - c_i)^{\beta-1}. \end{aligned} \tag{3.14}$$

Devido à complexidade da distribuição a posteriori (3.14), se faz necessário a utilização de métodos numéricos e, neste caso, para a estimação dos parâmetros dos itens e das proficiências foi utilizado o amostrador de Gibbs, algoritmo da classe dos Métodos de Monte Carlo via Cadeia de Markov (MCMC), implementado via software OpenBUGS (Bayesian inference Using Gibbs Sampling), direcionado para análise bayesiana de modelos complexos. O software é capaz de determinar um esquema MCMC apropriado, baseado nas distribuições condicionais completas.

3.3.6 Descrição das Distribuições Condicionais Completas

O amostrador de Gibbs fornece um esquema de gerações sucessivas a partir de distribuições conhecidas como distribuições condicionais completas permitindo a obtenção de amostras da distribuição a posteriori.

As distribuições condicionais completas para o ML3, indispensáveis para geração das amostras, são apresentadas em (3.15), (3.16), (3.17) e (3.18).

A distribuição condicional completa de a_i é dada por

$$\begin{aligned}
 p(a_i | \boldsymbol{\theta}, b_i, c_i, \mathbf{u}) &\propto \prod_{j=1}^n P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} h_1(a_i) \\
 &\propto \prod_{j=1}^n \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\
 &\times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}} \\
 &\times \left[\frac{1}{\sqrt{2\pi}\sigma_a} \exp \left\{ -\frac{1}{2} \left(\frac{a_i - \mu_a}{\sigma_a} \right)^2 \right\} I_{(0,\infty)}, i = 1, \dots, I. \right.
 \end{aligned} \tag{3.15}$$

A distribuição condicional completa de b_i é dada por

$$\begin{aligned}
 p(b_i | \boldsymbol{\theta}, a_i, c_i, \mathbf{u}) &\propto \prod_{j=1}^n P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} h_2(b_i) \\
 &\propto \prod_{j=1}^n \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\
 &\times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}} \\
 &\times \frac{1}{\sqrt{2\pi}\sigma_b} \exp \left\{ -\frac{1}{2} \left(\frac{b_i - \mu_b}{\sigma_b} \right)^2 \right\}, i = 1, \dots, I.
 \end{aligned} \tag{3.16}$$

A distribuição condicional completa de c_i é dada por

$$\begin{aligned}
 p(c_i | \boldsymbol{\theta}, a_i, b_i, \mathbf{u}) &\propto \prod_{j=1}^n P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} h_3(c_i) \\
 &\propto \prod_{j=1}^n \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\
 &\quad \times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}} \\
 &\quad \times \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} c_i^{\alpha-1} (1 - c_i)^{\beta-1}, i = 1, \dots, I.
 \end{aligned} \tag{3.17}$$

A distribuição condicional completa de θ_j é dada por

$$\begin{aligned}
 p(\theta_j | \boldsymbol{\zeta}, \mathbf{u}) &\propto \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}} g(\theta_j) \\
 &\propto \prod_{i=1}^I \left[c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right]^{u_{ji}} \\
 &\quad \times \left[1 - \left(c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \right) \right]^{1-u_{ji}} \\
 &\quad \times \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp \left\{ -\frac{1}{2} \left(\frac{\theta_j - \mu_\theta}{\sigma_\theta} \right)^2 \right\}, j = 1, \dots, n.
 \end{aligned} \tag{3.18}$$

3.3.7 Estimação das Proficiências

No Método de Estimação Bayesiano, as estimativas das proficiências são realizadas diretamente em conjunto com as estimações para os parâmetros dos itens, através da distribuição condicional completa para a variável latente θ_j dada em (3.18).

Uma das sugestões presentes na literatura para a estimação das proficiências é o método de estimação bayesiana marginal, a estimação das proficiências é feita em uma segunda etapa, assim como, na estimação por máxima verossimilhança marginal. De acordo com Andrade, Tavares e Valle (2000) supondo independência entre as proficiências de diferentes indivíduos, as estimações de suas proficiências podem ser feitas em separado.

Considerando a distribuição a priori dada em (3.9), para variável latente θ_j , com vetor de hiperparâmetros $\boldsymbol{\eta} = (\mu_\theta, \sigma_\theta^2)$ conhecidos, a densidade a posteriori para a proficiência do indivíduo j é dada por

$$p(\theta_j | \mathbf{u}_{j.}, \boldsymbol{\zeta}, \boldsymbol{\eta}) \propto L(\theta_j, \boldsymbol{\zeta}; \mathbf{u}_{j.}) g(\theta_j | \boldsymbol{\eta}), \tag{3.19}$$

onde $L(\theta_j, \boldsymbol{\zeta}; \mathbf{u}_{j.}) = \prod_{i=1}^I P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}$.

Assim, a estimação das proficiências dos indivíduos é baseada em alguma característica de $p(\theta_j | \mathbf{u}_{j.}, \boldsymbol{\zeta}, \boldsymbol{\eta})$, sendo as mais utilizadas a estimação pela média a posteriori (EAP) e a estimação pela moda a posteriori (MAP). Pode-se observar que a distribuição a posteriori dada em (3.19) é exatamente a distribuição condicional completa dada em (3.18).

Construção e Interpretação da Escala de Proficiência em Conteúdos Matemáticos

Para que seja possível construir uma escala de proficiência interpretável é conveniente que se tenha uma matriz de referência, na qual estejam estabelecidas competências (ou descritores) que abordem as habilidades que se deseja medir. É necessário o uso de um instrumento de medida para o traço latente investigado. Como mencionado anteriormente, na área educacional, são as provas que fazem o papel do instrumento de medida. Estas por sua vez, são compostas por itens (ou questões) elaboradas de acordo com as habilidades que se deseja aferir, dispostas na matriz de referência.

O presente trabalho, teve por objetivo a realização de um estudo sobre a construção de escalas de proficiência e a construção de uma escala de proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento de disciplinas de cálculo e similares, com base na TRI, bem como a sua interpretação pedagógica. O Público alvo da pesquisa foram os alunos ingressantes nos cursos das áreas de exatas da FCT/UNESP no ano letivo de 2017. Dando continuidade à pesquisa realizada por Rossi (2015).

A matriz de referência utilizada nesta pesquisa, foi a proposta por Rossi (2015) que, para elaboração da mesma, baseou-se nas matrizes de referência da disciplina de matemática, propostas para as avaliações aplicadas pelo SAEB (ANEAB, conhecida como prova do SAEB e a ANRESC, conhecida como Prova Brasil).

Para atingir os objetivos apresentados, se faz necessária a aplicação de uma prova, utilizada como instrumento de medida do conhecimento em questão. Os itens que compõem esta prova, estão baseados na prova elaborada por Rossi (2015) que por sua vez, baseou-se nos exemplos de itens propostos pelo Plano de Desenvolvimento da Educação (PDE) do SAEB e Prova Brasil (BRASIL, 2017f) e (BRASIL, 2017g). Entretanto, na atual proposta de prova, foram incluídos novos itens para algumas competências que haviam sido conjuntamente avaliadas em um mesmo item.

Neste capítulo será apresentada a matriz de referência utilizada neste trabalho bem como, a descrição dos itens que compõem a prova e como ocorreu a aplicação da mesma, a discussão dos resultados e a construção e interpretação da escala. Dessa forma, na Seção 4.1 é apresentada a matriz de referência utilizada, na Seção 4.2 discute-se a organização e descrição dos itens que fizeram parte da prova usada como instrumento de medida do conhecimento, na Seção 4.3 é abordado como se deu o processo de aplicação da prova,

na Seção 4.4 é apresentada a discussão dos resultados e na Seção 4.5 é apresentada a construção e interpretação da escala de proficiência em conteúdos matemáticos básicos.

4.1 Matriz de Referência

A matriz de referência adotada nesta pesquisa é composta em grande parte por competências pertencentes à matriz proposta para o 3º ano do EM utilizada na prova do SAEB somando 22 competências de um total de 36 que a compõem. Outras 10 competências fazem parte da matriz proposta para o 9º ano do EF utilizada na Prova Brasil e outras 4 competências foram acrescentadas a matriz, pois julgou-se serem necessárias para o propósito da prova, que é o de medir a proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo e similares.

A necessidade de acréscimo destas 4 competências, e a utilização de competências da matriz de referência para o 9º ano da Prova Brasil, fez com que fosse proposta essa nova matriz, ao invés de ter sido utilizada a matriz de referência para o 3º ano do EM proposta pelo SAEB.

É importante ressaltar aqui que cada competência selecionada para a matriz de referência proposta foi cuidadosamente analisada e as habilidades investigadas para cada uma delas, tem papel fundamental dentro dos conteúdos de cálculo. Por esta razão, a matriz de referência contempla inclusive competências pertencentes à matriz de referência do 9º ano utilizada na Prova Brasil. Apesar de, a princípio parecerem muito simples, a ausência dessas habilidades, faz com que os alunos não consigam prosseguir com sucesso nos conteúdos de cálculo e similares. Não se pode deixar de enfatizar também, que na elaboração da matriz de referência, Rossi (2015) contou com a colaboração de professores especialistas em cálculo do Departamento de Matemática e Computação e do programa de Pós-Graduação em Educação da FCT/UNESP.

Essa matriz de referência foi estruturada nos mesmos moldes das utilizadas como base e engloba três temas:

- Espaço e Forma;
- Grandezas e Medidas;
- Números e Operações/Álgebra e Funções.

Cada tema é composto por um conjunto de competências. Cada competência aborda uma habilidade que julgou-se ser necessária para o acompanhamento das disciplinas de cálculo e similares. Do mesmo modo que, as matrizes de referência dos sistemas de avaliação abordados no Capítulo 2 não contemplam todo currículo escolar, a matriz de referência proposta por Rossi (2015) também não abrange todo currículo escolar abordado no Ensino Médio. É composta por habilidades que espera-se que o aluno tenha desenvolvido ao final do Ensino médio, dentro do contexto proposto para este trabalho.

A matriz de referência com as competências separadas por tema pode ser consultada no Anexo 2, nos Quadros 2.1, 2.2 e 2.3, onde cada competência é caracterizada pela letra C e vem acompanhada da descrição da habilidade que se pretende avaliar e a informação sobre sua origem.

4.2 Elaboração do Instrumento de Avaliação

A seguir, serão apresentados os itens que compõem a prova proposta para ser aplicada aos alunos ingressantes no ano letivo de 2017 da FCT/UNESP nos cursos da área de exa-

tas. Como mencionado no início deste capítulo, a prova é composta em grande parte por itens que foram utilizados na prova aplicada por Rossi (2015), pois analisando os resultados obtidos pela mesma, estes itens se mostraram adequados para avaliar as habilidades para o qual foram propostos. Porém, como bem sugere Rossi (2015), alguns itens foram acrescentados e outros modificados a fim de estimar melhor as proficiências nos pontos localizados nas extremidades da escala, tanto inferior quanto superior. A relação desses itens é apresentada no Quadro 4.1.

Quadro 4.1: Relação dos itens que foram acrescentados e modificados para atual proposta de prova.

Competência	Prova Rossi (2015)	Prova Atual (2017)
C1	Avaliada pelo item 1 proposto para avaliar as competências C1 e C3.	Avaliada pelo item 1 incluído na prova para avaliar somente a competência C1.
C2	Avaliada pelo item 2 proposto para avaliar as competências C2 e C4.	O item foi modificado para avaliar somente a competência C2 e numerado como item 2 .
C3	Avaliada pelo item 1 proposto para avaliar as competências C1 e C3.	Este item foi mantido para avaliar somente a competência C3 e numerado como item 3.
C4	Avaliada pelo item 2 proposto para avaliar as competências C2 e C4.	Avaliada pelo item 4 incluído na prova para avaliar somente a competência C4.
C6	Avaliada pelo item 4 proposto para avaliar as competências C6 e C7.	O item foi modificado para avaliar somente a competência C6 e numerado como item 6.
C7	Avaliada pelo item 4 proposto para avaliar as competências C6 e C7.	Avaliada pelo item 7 incluído na prova para avaliar somente a competência C7.
C21	Avaliada pelo item 21 proposto para avaliar as competências C21 e C30.	Avaliada pelo item 21 incluído na prova para avaliar somente a competência C21.
C30	Avaliada pelo item 21 proposto para avaliar as competências C21 e C30.	O item foi modificado para avaliar somente a competência C30 e numerado como item 30.
C31	Avaliada pelo item 25.	O item foi modificado para avaliar a competência C31 e numerado como item 31.
C35	Avaliada pelo item 32.	Avaliada pelo item 35 incluído na prova para substituir o item proposto na prova anterior.

Fonte: Elaborado pela autora.

Assim, a prova elaborada para este trabalho contém um item para cada competência presente na matriz de referência, totalizando 36 itens. Estes estão dispostos a seguir, de acordo com o tema, com uma breve descrição e informações sobre a competência para a qual foi proposto e fonte de onde foi retirado. A alternativa destacada em negrito corresponde a alternativa correta e a prova no formato de aplicação pode ser consultada no Apêndice A.

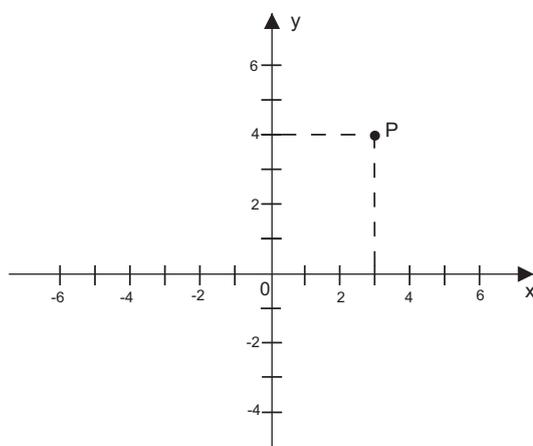
4.2.1 Tema I: Espaço e Forma

Este campo do conhecimento trata uma parte importante do currículo do ensino médio. As competências escolhidas para este tema, abordam habilidades de extrema importância para o estudo de funções.

C1 - Identificar a localização de pontos no plano cartesiano.

Esta competência pretende medir a habilidade de o aluno identificar adequadamente um ponto no plano a partir de seu par ordenado, ou vice-versa. O item escolhido como proposta foi retirado do PDE do SAEB como exemplo para o descritor 6 da matriz de referência para o 3º ano do Ensino Médio.

Item: A figura abaixo mostra um ponto em um plano cartesiano.



As coordenadas do ponto P são:

- (A) (6, 6).
- (B) (-3, 4).
- (C) (3, 4).**
- (D) (3, 7).
- (E) (4, 5).

C2 - Interpretar geometricamente os coeficientes da equação de uma reta.

Esta competência pretende avaliar a habilidade de o aluno reconhecer os coeficientes de uma equação de primeiro grau. O item proposto pode ser encontrado no PDE do SAEB como exemplo para o descritor 7 de sua matriz de referência.

Item: A reta da equação $2y + x = 0$:

- (A) é paralela ao eixo OX.
- (B) é paralela ao eixo OY.
- (C) tem coeficiente angular $-\frac{1}{2}$.**
- (D) tem coeficiente angular $\frac{1}{2}$.
- (E) tem coeficiente angular 2.

C3 - Identificar a equação de uma reta apresentada a partir de dois pontos dados ou de um ponto e sua inclinação.

Esta competência pretende avaliar a habilidade de o aluno construir a equação de uma reta a partir de dois de seus pontos ou então a partir de um ponto e de sua inclinação. O item proposto trata-se de um exemplo do PDE do SAEB indicado para o descritor 8 de sua matriz de referência.

Item: Qual é a equação da reta que contém os pontos $(3, 5)$ e $(4, -2)$?

(A) $y = -7x + 26$.

(B) $y = -\frac{1}{7}x - \frac{10}{7}$.

(C) $y = \frac{1}{7}x - \frac{18}{7}$.

(D) $y = x + 2$.

(E) $y = 7x - 16$.

C4 - Relacionar a determinação do ponto de interseção de duas ou mais retas com a resolução de um sistema de equações com duas incógnitas.

Esta competência tem a intenção de avaliar a habilidade de o aluno relacionar o conceito de resolução de problemas que envolvam um sistema de equações com duas incógnitas e o conceito de determinação do ponto de interseção de duas retas, dois importantes conceitos da matemática. Este item é baseado no exemplo para o descritor 9 da matriz de referência do PDE do SAEB.

Item: O ponto de interseção das retas de equações $x + 3y - 1 = 0$ e $x - y + 3 = 0$ é:

(A) $(1, -2)$.

(B) $(-2, 1)$.

(C) $(-1, -2)$.

(D) $(-2, -1)$.

(E) $(4, 1)$.

C5 - Reconhecer, dentre as equações do segundo grau com duas incógnitas, as que representam circunferências.

Esta competência tem por finalidade avaliar a capacidade de o aluno reconhecer a equação de uma circunferência, dentre um conjunto de equações do segundo grau. O item escolhido para esta competência é um exemplo encontrado no PDE do SAEB para o descritor 10 de sua matriz de referência.

Item: Dentre as equações abaixo, pode-se afirmar que a de uma circunferência é:

(A) $(x - 1)^2 + y^2 = 25$.

(B) $x^2 - y - 4x = -3$.

(C) $x^2 + y^2 = -16$.

(D) $x^2 - y - 9 = 0$.

(E) $x^2 - y^2 - 4x = 9$.

4.2.2 Tema II. Grandezas e Medidas

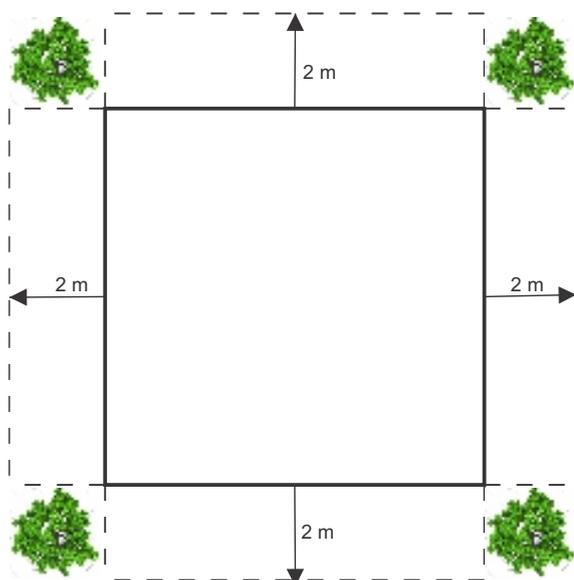
Para este tema foram selecionadas competências que buscam avaliar as habilidades dos alunos no cálculo do perímetro e área de figuras planas e área total e/ou volume de sólidos, sendo seu estudo de grande importância para o desenvolvimento da disciplina de

cálculo. Pode-se citar, por exemplo, as inúmeras aplicações da Integral no cálculo de áreas de regiões planas e volumes de sólidos.

C6 - Resolver problemas envolvendo cálculo de perímetro de figuras planas.

Esta competência pretende avaliar a habilidade de o aluno resolver problemas do cotidiano utilizando o cálculo de perímetro. O item proposto é um exemplo fornecido pelo PDE do SAEB para o descritor 11 de sua matriz de referência.

Item: Uma praça quadrada, que possui o perímetro de 24 metros, tem uma árvore próxima de cada vértice e fora dela. Deseja-se aumentar a área da praça, alterando-se sua forma e mantendo as árvores externas a ela, conforme ilustra a figura.



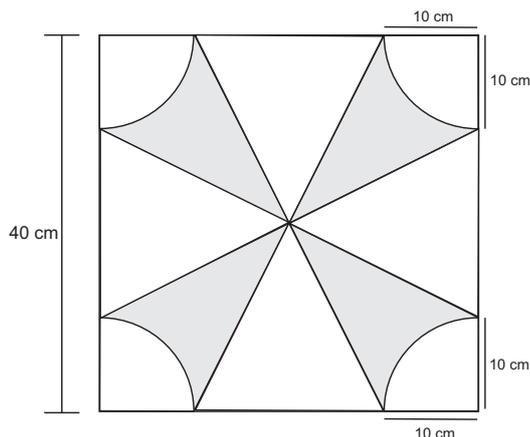
O novo perímetro da praça é:

- (A) 24 metros.
- (B) 32 metros.
- (C) 36 metros.
- (D) 40 metros.**
- (E) 64 metros.

C7 Resolver problemas envolvendo cálculo de área de figuras planas.

Pretende-se avaliar com essa competência a habilidade de o aluno trabalhar com cálculo de áreas envolvendo figuras planas. O item proposto é exemplo para o descritor 12 da matriz de referência do PDE do SAEB.

Item: Paulo resolve modificar o revestimento do piso de sua sala de estar e escolhe uma cerâmica cujo formato está representado na figura a seguir. A cerâmica escolhida tem a forma de um quadrado cujo o lado mede 40 cm e possui 4 arcos de circunferência, de raio igual a 10 cm, cujos centros estão localizados nos vértices do quadrado.



Com base nessas informações, qual é a área do desenho formado na cerâmica, em centímetros quadrados? (Considere $\pi = 3,14$)

- (A) 314.
- (B) 400.
- (C) **486.**
- (D) 1114.
- (E) 1286.

C8 - Resolver problemas envolvendo a área total e/ou volume de um sólido (prisma, pirâmide, cilindro, cone, esfera).

Esta competência pretende avaliar a habilidade de o aluno resolver problemas que envolvam cálculo de área de sólidos geométricos. O item proposto é baseado em um exemplo para o descritor 13 da matriz de referência do PDE do SAEB.

Item: Um corpo cilíndrico, com 4 cm de raio e 12 cm de altura, está com água até a altura de 8 cm. Foram colocadas em seu interior bolas de gude de 2 cm de diâmetro, e o nível da água atingiu a boca do vidro, sem derramamento.

Quantas bolas de gude foram colocadas?

- (A) 32.
- (B) **48.**
- (C) 64.
- (D) 80.
- (E) 96.

4.2.3 Tema III. Números e Operações/Álgebra e Funções

As competências selecionadas para este tema, abordam habilidades do conceito de número e suas operações. Espera-se que os alunos tenham desenvolvido a compreensão necessária e adequada de tal conceito ao fim do ensino médio, capacitando-os a fazer julgamentos matemáticos e a decidir quanto a estratégias de manipulação dos números e das operações. Conceito de extrema importância para o acompanhamento das disciplinas de cálculo e similares.

C9 - Identificar a localização de números reais na reta numérica.

O objetivo desta competência é avaliar a habilidade de o aluno representar a posição de números reais na reta numérica. O item proposto pode ser encontrado no PDE do SAEB como exemplo para o descritor 14 de sua matriz de referência.

Item: Na figura abaixo estão representados os números reais 0, x, y, 1.



A posição do produto $x \times y$ é:

(A) à esquerda do zero.

(B) entre 0 e x .

(C) entre x e y.

(D) entre y e 1.

(E) à direita de 1.

C10 - Resolver problemas com números naturais e inteiros envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão e potenciação).

Esta competência tem a finalidade de avaliar a habilidade de o aluno resolver problemas utilizando as operações indicadas com números naturais. O item proposto foi retirado do PDE da Prova Brasil para o ensino fundamental como exemplo para o descritor 19 da matriz de referência do 9º ano.

Item: Num cinema, há 12 fileiras com 16 poltronas e 15 fileiras com 18 poltronas. O número de poltronas é:

(A) 78.

(B) 192.

(C) 270.

(D) 462.

(E) 480.

C11 - Identificar a localização de números racionais na reta numérica.

O objetivo desta competência é avaliar a habilidade de o aluno localizar números racionais na reta. Além disso, reconhecer que entre dois números racionais existem infinitos outros racionais. Este item foi retirado do PDE da Prova Brasil como exemplo para o descritor 17 de sua matriz de referência.

Item: A figura abaixo mostra os pontos P e Q que correspondem a números racionais e foram posicionados na reta numerada do conjunto dos racionais.



Os valores atribuídos a P e Q, conforme suas posições na reta numérica são:

- (A) $P = -\frac{1}{5}$ e $Q = -\frac{3}{10}$.
 (B) $P = -\frac{3}{10}$ e $Q = -\frac{1}{5}$.
 (C) $P = -\frac{3}{5}$ e $Q = -\frac{7}{10}$.
 (D) $P = -\frac{7}{10}$ e $Q = -\frac{3}{5}$.
 (E) $P = -\frac{3}{5}$ e $Q = -\frac{1}{5}$.

C12 - Reconhecer diferentes representações de um número racional.

Esta competência pretende avaliar a habilidade de o aluno identificar números racionais nas suas diversas representações tais como fracionária, decimal ou percentual. O item foi retirado do PDE da Prova Brasil para o descritor 21 de sua matriz de referência.

Item: No Brasil $\frac{3}{4}$ da população vive na zona urbana. De que outra forma podemos representar esta fração?

- (A) 7,5%.
 (B) 15%.
 (C) 25%.
 (D) 34%.
 (E) **75%**.

C13 - Identificar fração como representação que pode estar associada a diferentes significados.

O objetivo desta competência consiste em avaliar a habilidade de o aluno reconhecer frações em diversas representações tais como partes de um inteiro, relação entre conjuntos e razão entre medidas. O item foi baseado no exemplo disponível no PDE da prova Brasil como sugestão para o descritor 22 de sua matriz de referência.

Item: Dos 11 jogadores de um time de futebol, apenas 5 tem menos de 25 anos de idade. A fração de jogadores com mais de 25 anos de idade é:

- (A) $\frac{5}{6}$.
 (B) $\frac{6}{5}$.
 (C) $\frac{5}{11}$.
 (D) $\frac{6}{11}$.
 (E) $\frac{5}{25}$.

C14 - Identificar frações equivalentes.

Esta competência pretende avaliar a habilidade de o aluno reconhecer que uma fração pode também ser representada por um conjunto infinito de outras frações equivalentes a ela. O item foi retirado do PDE da Prova Brasil como exemplo para o descritor 23 de sua

matriz de referência.

Item: Quatro amigos, João, Pedro, Ana e Maria saíram juntos para fazer um passeio por um mesmo caminho. Até agora, João andou $\frac{6}{8}$ do caminho; Pedro, $\frac{9}{12}$; Ana, $\frac{3}{8}$ e Maria, $\frac{4}{6}$. Os amigos que se encontrarão no mesmo ponto do caminho são:

(A) João e Pedro.

(B) João e Ana.

(C) Ana e Maria.

(D) Pedro e Ana.

(E) Maria e João.

C15 - Efetuar cálculos que envolvam operações (adição, subtração, multiplicação, divisão e potenciação) com números racionais.

Esta competência pretende avaliar a habilidade de o aluno efetuar cálculos de expressões com diferentes representações dos números racionais e envolvendo as operações básicas do conjunto dos racionais. O item foi baseado no exemplo para o descritor 25 da matriz de referência do 9º ano no PDE da prova Brasil.

Item: A professora de matemática propôs como exercício a expressão:

$$\left[\left(1 + \frac{1}{3} \right) \left(1 - \frac{1}{3} \right) \right]^2$$

Os alunos que resolveram corretamente a expressão encontraram como resultado:

(A) $-\frac{8}{9}$.

(B) 0.

(C) $\frac{8}{9}$.

(D) 2.

(E) $\frac{64}{81}$.

C16 - Resolver problemas com números racionais que envolvam as operações: adição, subtração, multiplicação, divisão e potenciação.

A finalidade desta competência consiste em avaliar a habilidade de o aluno resolver problemas utilizando operações com números racionais tais como adição, subtração, multiplicação, divisão e potenciação. O item foi retirado do PDE da Prova Brasil como exemplo para o descritor 26 de sua matriz de referência.

Item: Uma horta comunitária será criada em uma área de 5100 m². Para o cultivo de hortaliças, serão destinados $\frac{2}{3}$ dessa área. Quantos metros quadrados serão utilizados nesse cultivo?

(A) 170.

(B) 340.

(C) 1700.

(D) 2550.

(E) 3400.

C17 - Efetuar cálculos com valores aproximados de radicais.

Esta competência tem por objetivo avaliar a habilidade de o aluno resolver expressões com radicais não exatos, resolvendo os mesmos com aproximações. Este item foi retirado da prova elaborada por Rossi (2015) e pode ser consultado em (DANTE, 2000) p. 590.

Item: Supondo $\sqrt[4]{8} = 1,68$, o valor mais próximo de $\sqrt{\frac{0,09}{\sqrt{2}}}$ é:

- (A) 25, 2.
- (B) 0, 0252.
- (C) 0, 252.**
- (D) 2, 5.
- (E) 0, 00252.

C18 - Calcular o valor numérico de uma expressão algébrica.

Esta competência pretende avaliar a habilidade de o aluno substituir as variáveis da expressão por números inteiros e calcular seu valor numérico dada uma expressão algébrica envolvendo as várias operações. O item foi retirado do PDE da Prova Brasil como exemplo para o descritor 30 de sua matriz de referência.

Item: O resultado da expressão $2x^2 - 3x + 10$, para $x = -2$, é:

- (A) -4.
- (B) 0.
- (C) 12.
- (D) 13.
- (E) 24.**

C19 - Identificar a expressão algébrica que expressa uma regularidade observada em sequências de números ou figuras (padrões).

O item apresentado para esta competência pretende avaliar a habilidade de o aluno reconhecer a regularidade ocorrida em uma sequência e representá-la por meio de uma expressão algébrica. O item foi retirado do PDE da Prova Brasil como exemplo para o descritor 32 de sua matriz de referência.

Item: As variáveis n e P assumem valores conforme mostra o quadro abaixo.

n	5	6	7	8	9	10
P	8	10	12	14	16	18

A relação entre P e n é dada pela expressão:

- (A) $P = n + 1$.
- (B) $P = n + 2$.
- (C) $P = 2n - 2$.**
- (D) $P = n - 2$.
- (E) $P = 2n$.

C20 - Reconhecer expressão algébrica que representa uma função a partir de uma tabela.

O item sugerido para esta competência pretende avaliar a habilidade de o aluno identificar a expressão algébrica, que representa a função que governa os dados indicados em uma tabela. O item foi retirado do PDE do SAEB como exemplo de item para o descritor 18 de sua matriz de referência.

Item: Uma empresa, em processo de reestruturação, propôs a seus funcionários, admitidos há pelo menos dois anos, uma indenização financeira para os que pedissem demissão, que variava em função do número de anos trabalhados. A tabela abaixo era utilizada para calcular o valor (i) da indenização, em função do tempo trabalhado (t).

Tempo trabalhado (em anos)	Valor da indenização (em reais)
1	450
2	950
3	1450
4	1950

A expressão que permite determinar o valor de indenização i para t anos trabalhados é:

(A) $i = 450t$.

(B) $i = 450 + 500t$.

(C) $i = 450(t - 1)$.

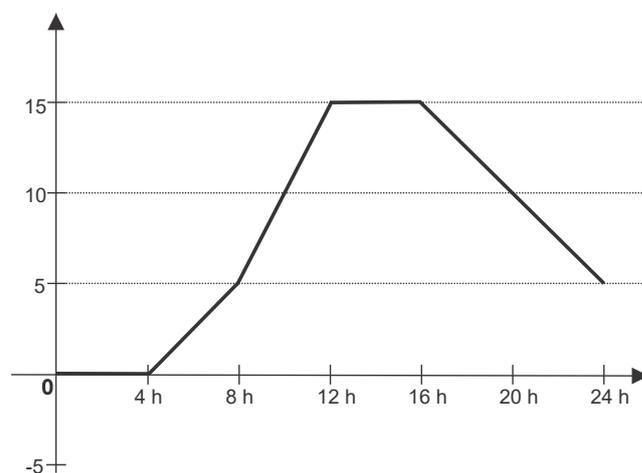
(D) $i = 450 + 500(t - 1)$.

(E) $i = 500t$.

C21 - Analisar crescimento/decrescimento, zeros de funções reais apresentadas em gráficos.

Com esta competência pretende-se avaliar a habilidade de o aluno identificar os zeros e/ou o crescimento e/ou decrescimento de qualquer função. O item proposto foi retirado do PDE do SAEB como exemplo para o descritor 20 de sua matriz de referência.

Item: O gráfico abaixo mostra a temperatura numa cidade da Região Sul, em um dia do mês de julho.



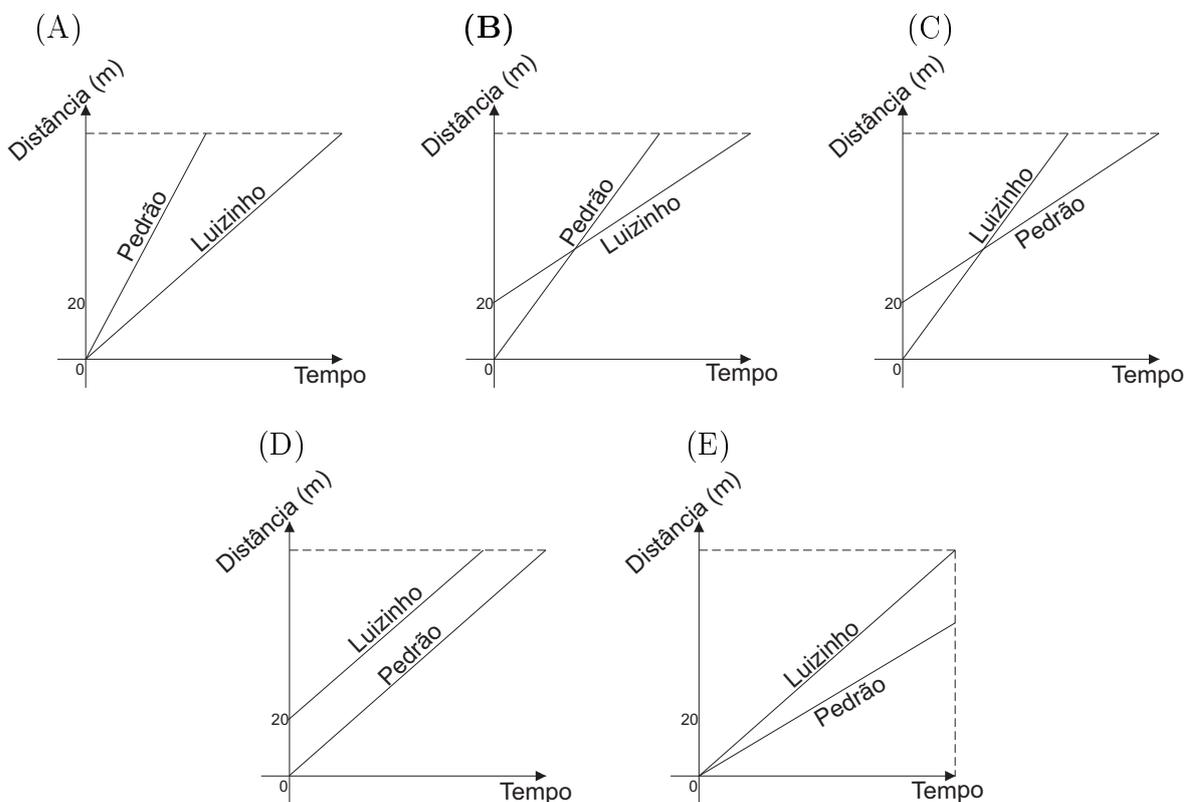
De acordo com o gráfico, a temperatura aumenta no período de:

- (A) 8 às 16h.
- (B) 16 às 24h.
- (C) 4 às 12h.
- (D) 12 às 16h.
- (E) 4 às 16h.

C22 - Identificar o gráfico que representa uma situação descrita no texto.

O objetivo desta competência é avaliar a habilidade de o aluno associar um gráfico à descrição de uma situação-problema. O item foi retirado do PDE do SAEB sugerido como exemplo para o descritor 21 de sua matriz de referência.

Item: Luizinho desafia seu irmão mais velho, Pedrão, para uma corrida. Pedrão aceita e permite que o desafiante saia 20 metros a sua frente. Pedrão ultrapassa Luizinho e ganha a corrida. O gráfico que melhor ilustra essa disputa é:



C23 - Identificar e resolver problemas que envolvam funções de primeiro grau.

Esta competência pretende avaliar a habilidade de o aluno resolver problemas do cotidiano que envolvam o conceito de funções, com o reconhecimento de regularidades numéricas ou geométricas, iniciado no ensino fundamental e ampliado no ensino médio. O item foi retirado do PDE do SAEB indicado para o descritor 19 de sua matriz de referência.

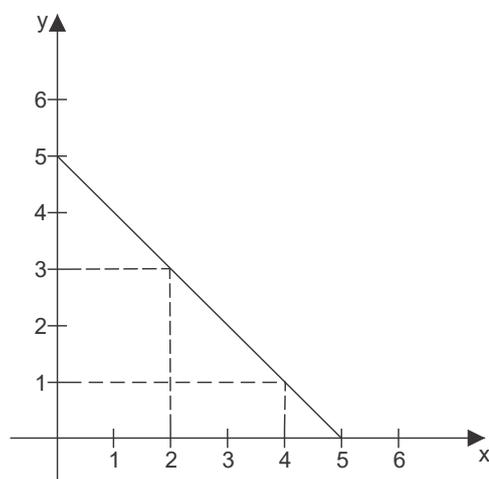
Item: O custo de produção de uma pequena empresa é composto por um valor fixo de R\$ 1.500,00 mais R\$ 10,00 por peça fabricada. O número x de peças fabricadas quando o custo é de R\$ 3.200,00 é:

- (A) 150.
- (B) 160.
- (C) 170.**
- (D) 320.
- (E) 470.

C24 - Reconhecer a representação algébrica de uma função do primeiro grau, dado o seu gráfico.

Com esta competência pretende-se avaliar a habilidade de o aluno associar a equação de uma função polinomial de primeiro grau ao seu gráfico. O item foi proposto com exemplo no PDE do SAEB para o descritor 24 de sua matriz de referência.

Item: O gráfico abaixo mostra uma reta em um plano cartesiano.



Qual é a equação da reta apresentada no gráfico?

- (A) $x - y - 5 = 0$.
- (B) $x + y - 5 = 0$.**
- (C) $x + y + 5 = 0$.
- (D) $x + y - 4 = 0$.
- (E) $x + y = 6$.

C25 - Identificar e resolver problemas de inequações de primeiro grau.

Esta competência tem o objetivo de avaliar a habilidade de o aluno resolver situações problema que envolvam o conceito de inequações. O item foi retirado da prova elaborada por Rossi (2015) e pode ser encontrado em (DANTE, 2000) p. 64.

Item: Por uma mensagem dos Estados Unidos para o Brasil, via fax, a Empresa de Correios e Telégrafos (ECT) cobra R\$ 1,37 pela primeira página e R\$ 0,67 por página que se segue, completa ou não. Qual o número mínimo de páginas de uma dessas mensagens para que seu preço ultrapasse o valor de R\$ 10,00?

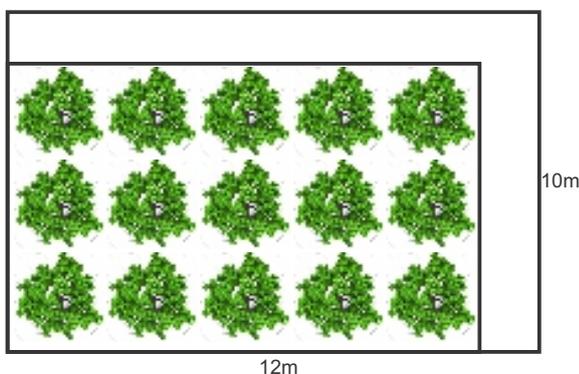
- (A) 8.
- (B) 10.

- (C) 12.
(D) 14.
 (E) 16.

C26 - Identificar e resolver problemas que envolvam equações de segundo grau.

O objetivo desta competência é avaliar a habilidade de o aluno resolver problemas em que seja necessário utilizar uma equação de segundo grau. o item foi retirado do PDE do SAEB como exemplo para o descritor 17 de sua matriz de referência.

Item: Em um terreno retangular de tamanho $10\text{m} \times 12\text{m}$, deseja-se construir um jardim com 80 m^2 de área, deixando uma faixa para o caminho (sempre de mesma largura), como mostra a figura.



A largura do caminho deve ser de:

- (A) 1 metro.
 (B) 1,5 metros.
(C) 2 metros.
 (D) 2,5 metros.
 (E) 3 metros.

C27 - Identificar e resolver problemas de inequações de segundo grau.

Esta competência pretende avaliar a habilidade de o aluno resolver problemas que envolvam a utilização de inequações de segundo grau. O item foi retirado da prova proposta por Rossi (2015) e pode ser encontrado em (DANTE, 2000) p. 594.

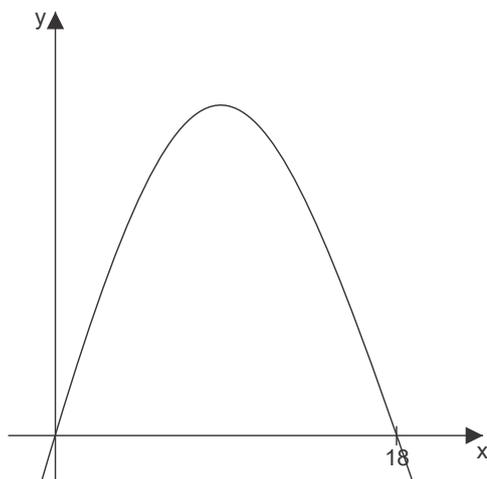
Item: Seja M o conjunto dos números naturais n , tal que $2n^2 - 75n + 700 \leq 0$. Assim, é correto afirmar que:

- (A) apenas um dos elementos de M é múltiplo de 4.**
 (B) apenas dois dos elementos de M são primos.
 (C) a soma de todos os elementos de M é igual a 79.
 (D) M contém exatamente 6 elementos.
 (E) os elementos de M são múltiplos de 3.

C28 - Resolver problemas que envolvam os pontos de máximo ou de mínimo no gráfico de uma função polinomial de segundo grau.

Esta competência tem por objetivo avaliar a habilidade de o aluno resolver problemas relacionados com os pontos de máximo ou de mínimo de uma função polinomial de segundo grau. O item foi retirado da prova elaborada por Rossi (2015).

Item: Uma bala é atirada de um canhão e sua trajetória descreve uma parábola de equação $y = -5x^2 + 90x$, onde as variáveis x e y são medidas em metros.



Nessas condições, a altura máxima atingida pela bala é:

- (A) 30,0 metros.
- (B) 40,5 metros.
- (C) 81,5 metros.
- (D) 405 metros.**
- (E) 810 metros.

C29 - Identificar e resolver equações modulares.

O objetivo desta competência é avaliar a habilidade de o aluno identificar e resolver equações modulares. O item foi retirado da prova elaborada por Rossi (2015) e pode ser encontrado em (DANTE, 2000) p.95.

Item: O conjunto solução da equação $|3x - 5| = 5x - 1$ é:

- (A) $\{-2\}$.
- (B) $\{\frac{3}{4}\}$.**
- (C) $\{\frac{1}{5}\}$.
- (D) $\{2\}$.
- (E) $\{\frac{3}{4}, -2\}$.

C30 - Identificar e resolver inequações modulares.

Esta competência pretende avaliar a habilidade de o aluno identificar e resolver inequações modulares. Este item foi baseado no proposto para essa competência na prova elaborada por Rossi (2015).

Item: São dados os conjuntos $A = \{x \in \mathbb{Z} \mid |x - 1| < 3\}$ e $B = \{x \in \mathbb{Q} \mid |x| < 4\}$.

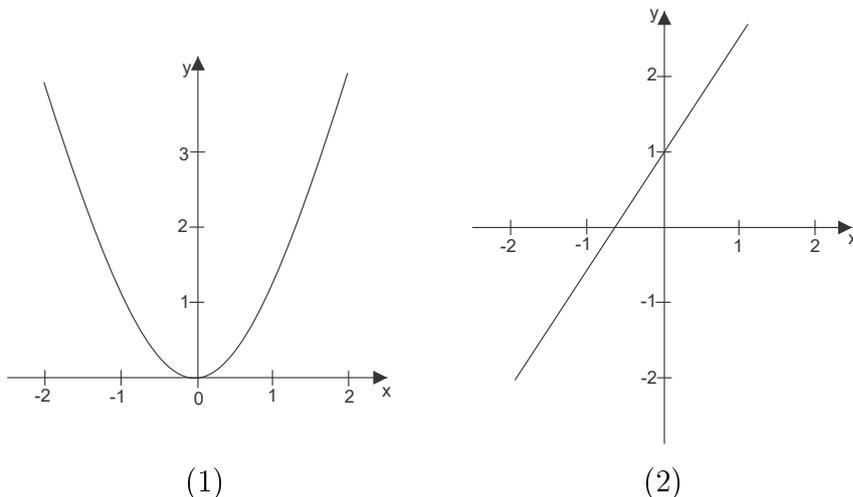
É correto afirmar que:

- (A) $A \cap B = \emptyset$.
- (B) $A \cup B = B$.
- (C) $A \subset B$ e $A \neq B$.
- (D) $A \cap B$ tem dois elementos.
- (E) $A \cup B$ tem infinitos elementos.**

C31 - Reconhecer o gráfico de uma função polinomial de primeiro grau por meio de seus coeficientes.

Esta competência tem o objetivo de avaliar a habilidade de o aluno manusear os coeficientes linear e angular da reta de forma a identificar o gráfico de uma função polinomial de primeiro grau. Este item está baseado no proposto para essa competência na prova elaborada por Rossi (2015).

Item: Observe os gráficos



Qual dos gráficos representa uma função polinomial de primeiro grau?

- (A) o gráfico (2).**
- (B) o gráfico (1).
- (C) os dois gráficos.
- (D) o gráfico (2) apenas considerando a parte positiva.
- (E) Nenhum dos gráficos.

C32 - Relacionar as raízes de um polinômio com sua decomposição em fatores do primeiro grau.

Esta competência pretende avaliar a habilidade de o aluno decompor um polinômio em fatores do primeiro grau. O item está baseado no exemplo do PDE do SAEB sugerido para o descritor 26 de sua matriz de referência.

Item: As raízes do polinômio $P(x) = (x - 3)^2(x + 1)$ são:

- (A) -2 e 1 .
- (B) 3 e -1 .**
- (C) -3 e 1 .
- (D) 3 e 1 .

(E) -3 e -1 .

C33 - Identificar a representação gráfica e/ou algébrica de uma função exponencial.

Esta competência pretende avaliar a habilidade de o aluno identificar a representação algébrica ou gráfica de uma função exponencial. O item faz parte do PDE do SAEB, e foi proposto como exemplo para avaliar o descritor 27 da sua matriz de referência.

Item: Abaixo estão relacionadas algumas funções. Entre elas, a função exponencial crescente é:

(A) $f(x) = 5^{-x}$

(B) $f(x) = \left(\frac{\sqrt{3}}{2}\right)^2$

(C) $f(x) = (0,1)^x$

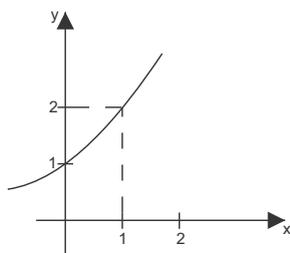
(D) $f(x) = 10^x$

(E) $f(x) = 0,5^x$

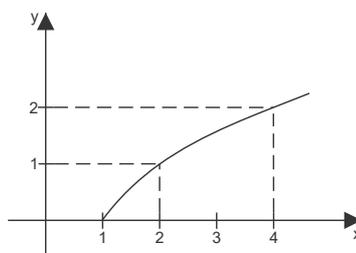
C34 - Identificar a representação gráfica e/ou algébrica de uma função logarítmica, reconhecendo-a como inversa da função exponencial.

O objetivo desta competência consiste em avaliar a habilidade de o aluno reconhecer a representação algébrica ou gráfica de uma função logarítmica e associá-la a uma função exponencial. O item foi retirado do PDE do SAEB proposto para avaliar o descritor 28 de sua matriz de referência.

Item: Nos gráficos (1) e (2) abaixo estão representadas duas funções.



(1)



(2)

Pode-se afirmar que:

(A) $y = 2x$ está representada no (1).

(B) $y = x^2 + 1$ está representada no (2).

(C) $y = \log_2 x$ está representada no (2).

(D) $y = 2^x$ está representada no (2).

(E) $y = \log x$ está representada no (2).

C35 - Resolver problemas que envolvam função exponencial.

Esta competência é utilizada para avaliar a habilidade de o aluno resolver problemas envolvendo a função exponencial. Muito comum em contextos que abordam fenômenos tais como químicos ou biológicos. Este item foi retirado da prova elaborada por Esteves (2015).

Item: A população de bactérias em um meio pode ser modelada, sob certas condições, por uma função exponencial. Sabendo que a cada hora esta população duplica e que no tempo $t = 0$ existem 100 bactérias, a função $P(t)$ que representa a população em função do tempo t (em horas) é:

(A) $p(t) = 100 \times 2^t$.

(B) $p(t) = 100 + 2t$.

(C) $p(t) = 2^t + 100$.

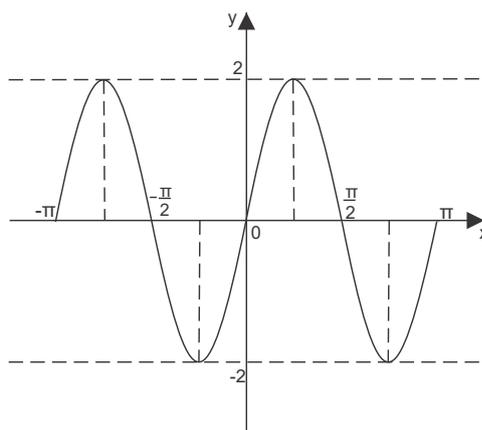
(D) $p(t) = 2t$.

(E) $p(t) = 2^t$.

C36 - Identificar a representação gráfica e/ou algébrica de funções trigonométricas (seno, cosseno, tangente), reconhecendo suas propriedades.

Esta competência pretende avaliar a habilidade de o aluno identificar o gráfico que representa uma função trigonométrica, bem como, dado um gráfico, identificar a função trigonométrica que o representa. Este item foi retirado da prova elaborada por Rossi (2015) e pode ser consultado em (DANTE, 2000) p. 224.

Item: O gráfico abaixo representa a função:



(A) $y = -2 \cos x$.

(B) $y = \cos \frac{x}{2}$.

(C) $y = 2 \operatorname{sen} x$.

(D) $y = \operatorname{sen} \frac{x}{2}$.

(E) $y = 2 \operatorname{sen} 2x$.

4.3 Aplicação da Avaliação

As provas foram aplicadas na FCT/UNESP, para os alunos matriculados no primeiro semestre do primeiro ano, na disciplina de Cálculo Diferencial e Integral I e Matemática Elementar, dos cursos da área de exatas sendo estes:

- Ciência da Computação;
- Engenharia Ambiental;
- Engenharia Cartográfica;
- Estatística;
- Física;
- Matemática Diurno;
- Matemática Noturno;
- Química.

A aplicação das provas foi realizada conforme cada docente responsável pela disciplina de Cálculo Diferencial e Integral I disponibilizou o seu horário de aula e, portanto, não foram realizadas no mesmo dia. Exceto para os cursos de Matemática diurno e noturno, para os quais foram utilizadas as aulas de Matemática elementar, pois a disciplina de cálculo nesses cursos só é oferecida a partir do segundo semestre do primeiro ano. Durante todo o processo de aplicação da prova, e mesmo após o término das aplicações em todas as turmas, não foi permitido em nenhum momento que os alunos ficassem em posse do caderno de questões, fizessem cópias ou fotografassem as mesmas a fim de evitar a exposição dos itens.

Todas as provas foram aplicadas nos meses de julho e agosto, ainda no primeiro semestre do ano letivo de 2017, que se encerrou na FTC/UNESP em 23 de agosto de 2017. A primeira Aplicação ocorreu em 25 de julho no curso de Engenharia Ambiental com a participação de 28 respondentes, no dia 26 de julho foi aplicada a prova para o curso de Física com 38 respondentes, no dia 28 de julho foram feitas as aplicações da prova no curso de Ciência da Computação com 32 respondentes e no curso de Química com 28 respondentes. No dia 31 de julho foram aplicadas provas para os cursos de Matemática diurno com 20 respondentes e Matemática noturno com 38 respondentes. Em 14 de agosto foi feita a aplicação da prova no curso de Estatística com a participação de 36 respondentes e, por fim, no dia 15 de agosto foi aplicada a prova no curso de Engenharia Cartográfica, com a participação de 57 respondentes, totalizando 277 respondentes.

Para a aplicação da prova, foram entregues aos alunos um caderno contendo 36 questões de múltipla escolha, cada uma com cinco alternativas, contendo somente uma correta, duas cópias de um Termo de Consentimento Livre e Esclarecido, para que o aluno pudesse assinar uma das cópias, autorizando a utilização de suas respostas para a pesquisa, ficando em poder da outra cópia para qualquer tipo de contato futuro e um cartão de respostas para os alunos transcreverem suas respostas a fim de facilitar a leitura das mesmas. Todos os documentos citados podem ser consultados nos Apêndices A, B e Anexo 1 respectivamente.

Ao final da prova os alunos participantes fizeram a entrega do caderno de questões, do cartão resposta devidamente preenchido e uma cópia assinada do Termo de Consentimento Livre e Esclarecido.

4.4 Discussão dos Resultados

Terminada a etapa de aplicação das provas para os alunos matriculados no primeiro semestre do primeiro ano, na disciplina de Cálculo Diferencial e Integral I e Matemática Elementar, dos cursos da área de exatas da FCT/UNESP, como descrito na seção anterior, foi dado início ao processo de leitura dos cartões resposta. Para que este processo fosse realizado com sucesso, contou-se com a participação de um professor do Departamento de Engenharia Cartográfica da FCT/UNESP, que conduziu a leitura, utilizando uma leitora pertencente ao seu departamento. As leituras dos cartões efetuadas pela leitora foram transferidas para um arquivo txt. Feito isto, todos os cartões resposta foram conferidos com a leitura efetuada na leitora para identificar possíveis falhas e garantir a sequência dos cartões de acordo com a sequência da leitura. Desta forma, pode-se identificar cada um dos respondentes caso seja necessário.

Após a conferência da leitura dos cartões utilizou-se a linguagem de programação R (TEAM, 2015) para fazer a leitura do arquivo txt com as respostas, conferi-las de acordo com o gabarito da prova e transformá-las para forma dicotômica, ou seja, apresentando 1 para respostas corretas, 0 para respostas incorretas. As respostas em branco ou anuladas foram consideradas incorretas. Foram utilizadas as funções *recode.variables* do pacote *Deducer* para converter as respostas dadas em letras (A, B, C, D ou E) para números (1, 2, 3, 4 ou 5 respectivamente) e *mult.choice* do pacote *ltm* para transformar as respostas em 0 ou 1 de acordo com o gabarito.

Para compor a discussão dos resultados, a Subseção 4.4.1 traz uma análise descritiva dos dados, realizada através da linguagem de programação R.

A Subseção 4.4.2 apresenta uma análise dos resultados obtidos com base na TRI. Para obter as estimativas dos parâmetros dos itens e também da proficiência dos respondentes com o uso da TRI, foi aplicado o modelo logístico unidimensional de três parâmetros especificado na Seção 3.2. Para este processo de calibração dos itens e estimação da proficiência dos respondentes utilizou-se o software OpenBUGS direcionado para inferência bayesiana em modelos complexos, que faz uso do Amostrador de Gibbs para a obtenção das estimativas para os parâmetros de interesse. Os gráficos e tabelas apresentados foram construídos através da linguagem de programação R. As estimativas para os parâmetros dos itens e das proficiências foram geradas a partir de uma amostra de tamanho 40 mil, descartadas (*burn-in*) as primeiras 5000, iterações para tirar o efeito dos valores iniciais, e um espaçamento entre os pontos amostrados (*thin*) de 10, para diminuir a correlação dentro da cadeia. A convergência da cadeia foi observada pela sua trajetória para cada um dos parâmetros ao longo das iterações e utilizando o software R, foi feito o teste de *Heidberger and Welch* do pacote CODA que indicou a convergência de todos os parâmetros dos itens.

Foi utilizada a versão 3.2.3 do software OpenBUGS. O tempo utilizado para estimação dos parâmetros dos itens e das proficiências foi de 1743 segundos. Os programas foram executados em computador com processador Intel(R) Core(TM) i3 - 5005U CPU @ 2.00 GHz 2.00 GHz.

Os parâmetros dos itens e das proficiências foram gerados a partir da especificação das seguintes distribuições a priori com seus respectivos hiperparâmetros:

- $a_i \sim Normal(0, 1)I(0,)$, $i = 1, \dots, I$ sendo, 0 e 1 os valores da média e da variância respectivamente;
- $b_i \sim Normal(0, 2)$, $i = 1, \dots, I$ sendo, 0 e 2 os valores da média e da variância respectivamente;

- $c_i \sim \text{Beta}(2, 8)$, $i = 1, \dots, I$ sendo, $\alpha = 2$ e $\beta = 8$, levando para média a priori 0,2 e variância 0,015;
- $\theta_j \sim \text{Normal}(0, 1)$, $j = 1, \dots, n$ sendo, 0 e 1 os valores de média e variância respectivamente, uma vez que as estimativas da proficiência dos indivíduos são inicialmente obtidas na escala da distribuição normal padrão e, posteriormente, para a construção e interpretação da escala, transformada conforme descrito na Seção 2.2 para obter os valores na escala (250, 50).

Os hiperparâmetros para a distribuição do parâmetro a_i foram definidos a partir de testes de sensibilidade, buscando atribuir valores que permitissem a especificação de priors pouco informativas. Tanto para o parâmetro a_i quanto para o parâmetro b_i foram consideradas sugestões da literatura.

A escolha para a distribuição dos c_i são subjetivas e baseadas na ideia de haver cinco alternativas para cada item, sendo somente uma correta.

Os resultados foram comparados aos obtidos na estimação dos parâmetros desses mesmos itens através do programa BILOG-MG. Observou-se nesta comparação que os resultados bayesianos e clássicos são similares, reforçando a opção por distribuições a priori pouco informativas. O programa BILOG-MG é um software amplamente utilizado em testes educacionais para estimação dos parâmetros dos itens e proficiências com base na TRI. Como citado anteriormente, é o programa utilizado para estimação dos parâmetros dos itens e proficiências nas provas do SAEB, ENEM e SARESP.

4.4.1 Análise Descritiva dos Resultados

A Tabela 4.1 apresenta os resultados descritivos para o mínimo, média e máximo de acertos por sala de cada curso avaliado da FCT/UNESP. Foram considerados inicialmente somente 35 itens dos 36 que compuseram a prova, pois após a análise dos itens foi constatado que o item 30, modificado para a prova atual, possuía três alternativas corretas e por este motivo não foi considerado nas análises apresentadas a seguir.

Tabela 4.1: Número mínimo, média e máximo de acertos por sala para cada curso avaliado de um total de 35 itens considerados.

Curso	Mínimo	Média	Máximo	Nº de Respondentes
Ciência da Computação	16	25,56	35	32
Engenharia Ambiental	9	22,25	33	28
Engenharia Cartográfica	10	19,58	29	57
Estatística	8	22,44	30	36
Física	10	19,79	31	38
Matemática Diurno	13	20,80	31	20
Matemática Noturno	8	19,29	33	38
Química	9	18,07	30	28

Fonte: Elaborado pela autora.

A Tabela 4.2 também traz como resultado descritivo para cada item, considerado o número de acertos, o número de respostas ausentes ou anuladas (NA's), o total de respostas obtidas desconsiderando-se os NA's para um total de 277 alunos respondentes e a porcentagem de acerto. A Tabela 4.2 está organizada em ordem decrescente de acordo com o percentual de respostas corretas para cada item. Pode-se observar que o item 1 obteve o maior número de respostas corretas, 274 de um total de 277 respostas apresentadas para

este item (98,92%). Já o item 36 foi o item que recebeu o menor número de respostas corretas, 27 respostas corretas de um total de 275 apresentadas para o item (9,82%).

Tabela 4.2: Número de acertos, de respostas ausentes, total de respostas e percentual de acerto para cada item considerado de 277 respondentes.

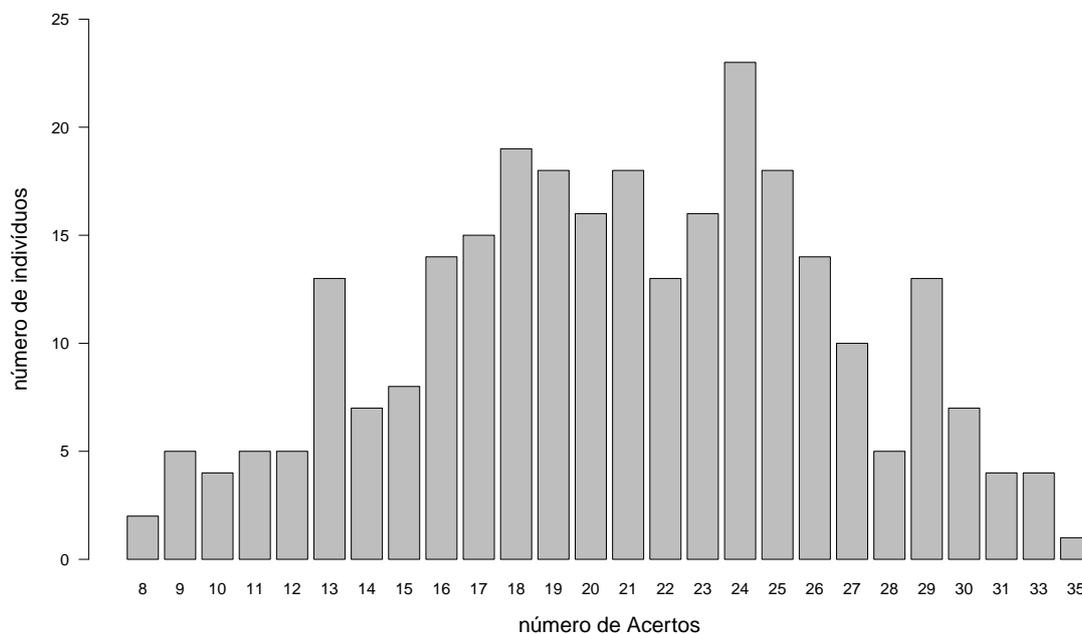
Item	Total de Acertos	Respostas Ausentes	Total de Respostas	%
1	274	0	277	98,92
10	262	0	277	94,58
12	252	0	277	90,97
19	243	0	277	87,73
13	236	0	277	85,20
23	235	0	277	84,84
21	230	0	277	83,03
16	226	1	276	81,88
18	226	0	277	81,59
14	206	0	277	74,37
33	203	0	277	73,29
20	201	0	277	72,56
26	191	0	277	68,95
22	187	0	277	67,51
15	185	1	276	67,03
24	174	0	277	62,82
11	173	1	276	62,68
4	171	1	276	61,96
3	161	0	277	58,12
31	160	0	277	57,76
25	152	0	277	54,87
6	144	0	277	51,99
9	141	0	277	50,90
35	141	0	277	50,90
32	135	2	275	49,09
7	130	2	275	47,27
28	128	4	273	46,89
2	116	2	275	42,18
34	97	0	277	35,02
5	83	2	275	30,18
27	79	3	274	28,83
8	76	2	275	27,64
17	69	4	273	25,27
29	58	0	277	20,94
36	27	2	275	9,82

Fonte: Elaborado pela autora.

Também como resultado descritivo é apresentado na Figura 4.1 um gráfico de barras que relaciona o número de acertos com o número de indivíduos. Pode-se observar que apenas um indivíduo obteve 35 acertos, para o restante dos valores apresentados, tem-se mais de um indivíduo com o mesmo número de acertos. Isto porém, não significa que os indivíduos que acertaram o mesmo número de questões possuem a mesma proficiência. O cálculo da proficiência a partir da TRI leva em consideração outros aspectos, tais como

os parâmetros dos itens e o padrão de resposta do indivíduo e não apenas a quantidade de acertos. Desta maneira, a proficiência de dois indivíduos com a mesma quantidade de acertos pode ser diferente, pois depende de quais itens estão certos e errados.

Figura 4.1: Relação entre o número de acertos e número de indivíduos.



Fonte: Elaborado pela autora.

4.4.2 Análise dos Resultados Obtidos com Base na Teoria da Resposta ao Item

As estimativas referentes aos parâmetros dos itens e também das proficiências foram obtidas com o uso do software OpenBUGS, como citado no início deste capítulo. Com a análise inicial dos resultados, optou-se por excluir também o item 1, respondido corretamente por quase todos os participantes, 274 dos 277 respondentes, considerado um item muito fácil. Este item mostrou pouco poder de discriminação e coeficiente de correlação bisserial negativo, análise feita no programa BILOG- MG. Na Tabela 4.3 podem ser observadas as estimativas obtidas para os parâmetros de discriminação (a_i), de dificuldade (b_i) e de acerto ao acaso (c_i) para os 34 itens considerados, bem como seus respectivos intervalos de 95% de credibilidade. Os valores obtidos para o parâmetro a_i foram satisfatórios, todos bem próximos ou maiores que 1, exceto para o item 25 que possui o menor valor para o parâmetro a , estimado em 0,3803, indicando que este item tem baixo poder de discriminação entre os indivíduos. Embora apresente baixo poder de discriminação, apresentou coeficiente de correlação bisserial positivo e valor de discriminação próximo de 0,4. Já os itens 04 e 36 possuem o maior poder de discriminação, com valor do parâmetro a estimado em 1,818 e 1,825 respectivamente.

Em relação à dificuldade dos itens, a maior parte dos valores se encontram no intervalo de -3 a 3 , intervalo onde, em geral, se concentram a maioria dos valores para o parâmetro. Pode-se observar que o item com o menor valor para o parâmetro de dificuldade ($-2,553$) é o 10, sendo este considerado o mais fácil dos itens. Por outro lado, o item 5 foi considerado o mais difícil, com o maior valor para o parâmetro de dificuldade ($2,754$).

O valor esperado para o parâmetro de acerto ao acaso deve estar em torno de 0,2, considerando que cada um dos itens que compõem a prova possuem 5 alternativas e apenas uma correta. Pode ser observado na Tabela 4.3 que os itens 4, 21, 31, 32 e 33 possuem probabilidade de acerto ao acaso em torno de 0,25. Este fato pode indicar que alguma alternativa desses itens é desprezada pelo respondente na hora de escolher a resposta correta, aumentando assim a probabilidade de acerto ao acaso das alternativas restantes. O item 4 tem a maior probabilidade de acerto ao acaso (0,2657) e o item 36 possui a menor probabilidade de acerto ao acaso (0,0462).

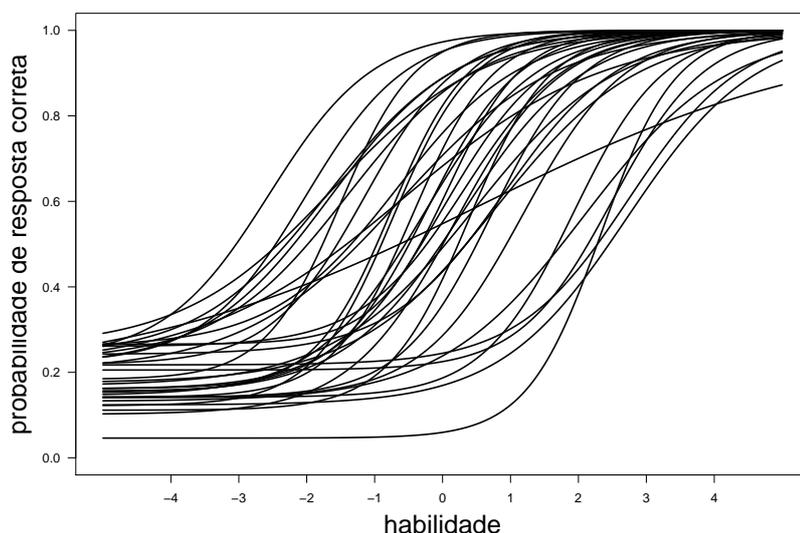
Tabela 4.3: Estimativas para os parâmetros a_i de discriminação, b_i de dificuldade e c_i de acerto ao acaso e seus respectivos intervalos de credibilidade (IC) de 95%.

Item	\hat{a}_i	IC	\hat{b}_i	IC	\hat{c}_i	IC
2	1,49600	[0,83670 ; 2,55000]	0,72680	[0,32780 ; 1,1500]	0,14180	[0,02925 ; 0,26020]
3	1,28500	[0,79120 ; 2,03800]	-0,00850	[-0,47250 ; 0,5110]	0,16120	[0,02844 ; 0,33450]
4	1,81800	[0,95660 ; 3,03200]	0,08940	[-0,45180 ; 0,5212]	0,26570	[0,07930 ; 0,41570]
5	1,20600	[0,33360 ; 2,62400]	2,75400	[1,82300 ; 4,1660]	0,21740	[0,04753 ; 0,31440]
6	1,13200	[0,66130 ; 1,84100]	0,30040	[-0,20000 ; 0,8530]	0,15130	[0,02504 ; 0,31570]
7	0,95690	[0,44940 ; 1,82400]	0,79820	[0,13840 ; 1,5620]	0,17550	[0,03178 ; 0,34280]
8	1,46700	[0,44120 ; 2,87100]	2,50200	[1,74200 ; 3,7260]	0,20540	[0,07119 ; 0,28640]
9	1,29000	[0,84990 ; 1,82500]	0,18030	[-0,15780 ; 0,5720]	0,10170	[0,01690 ; 0,22870]
10	1,29600	[0,77980 ; 1,97500]	-2,55300	[-3,63100 ; -1,7200]	0,23340	[0,03271 ; 0,53380]
11	1,23300	[0,78030 ; 1,83300]	-0,26950	[-0,69990 ; 0,2179]	0,14600	[0,02108 ; 0,31060]
12	1,33900	[0,82930 ; 1,99200]	-2,02100	[-2,84100 ; -1,3180]	0,22130	[0,03428 ; 0,50350]
13	1,04900	[0,63030 ; 1,60300]	-1,68400	[-2,67800 ; -0,8086]	0,23830	[0,03922 ; 0,52110]
14	1,76500	[1,17800 ; 2,55400]	-0,75550	[-1,14800 ; -0,3545]	0,15650	[0,02926 ; 0,33730]
15	1,69200	[1,15400 ; 2,42600]	-0,43550	[-0,80090 ; -0,0571]	0,14170	[0,02132 ; 0,31200]
16	1,44600	[0,89740 ; 2,14200]	-1,15300	[-1,80300 ; -0,5252]	0,21860	[0,03341 ; 0,46920]
17	1,50600	[0,64420 ; 2,78800]	1,91900	[1,38300 ; 2,7310]	0,14170	[0,04770 ; 0,22000]
18	1,04600	[0,60910 ; 1,59300]	-1,39900	[-2,37400 ; -0,5261]	0,22900	[0,03600 ; 0,49590]
19	1,80000	[1,14800 ; 2,60300]	-1,51100	[-2,05300 ; -1,0290]	0,18350	[0,03074 ; 0,41650]
20	1,69500	[1,19700 ; 2,32600]	-0,73400	[-1,07300 ; -0,3828]	0,12200	[0,01755 ; 0,28150]
21	0,90310	[0,50640 ; 1,44700]	-1,61400	[-2,78500 ; -0,4926]	0,25800	[0,03671 ; 0,55520]
22	0,67130	[0,31740 ; 1,13500]	-0,58280	[-1,67000 ; 0,6112]	0,21350	[0,03616 ; 0,45220]
23	1,07900	[0,64140 ; 1,62300]	-1,66700	[-2,64600 ; -0,8875]	0,21580	[0,03287 ; 0,47290]
24	1,48300	[0,93470 ; 2,20100]	-0,22880	[-0,65160 ; 0,2051]	0,16170	[0,03033 ; 0,32670]
25	0,38030	[0,05865 ; 1,02000]	0,55930	[-1,08500 ; 2,3960]	0,18230	[0,03040 ; 0,43620]
26	0,86730	[0,47210 ; 1,40400]	-0,62760	[-1,48600 ; 0,2729]	0,20400	[0,03051 ; 0,43700]
27	0,94600	[0,43440 ; 1,84900]	2,08300	[1,34500 ; 3,1650]	0,13230	[0,02423 ; 0,24490]
28	1,67600	[1,07300 ; 2,49800]	0,38140	[0,08397 ; 0,7155]	0,11140	[0,02370 ; 0,21800]
29	1,06800	[0,46950 ; 2,06700]	2,71700	[1,85300 ; 4,0130]	0,12370	[0,03404 ; 0,20500]
31	1,31200	[0,60740 ; 2,48000]	0,32820	[-0,36680 ; 0,9240]	0,26060	[0,06184 ; 0,43040]
32	1,17400	[0,43650 ; 2,33400]	0,91170	[0,17950 ; 1,6590]	0,24220	[0,05517 ; 0,39040]
33	1,09000	[0,58790 ; 1,87600]	-0,67380	[-1,54500 ; 0,1823]	0,25970	[0,04141 ; 0,50060]
34	1,34200	[0,70770 ; 2,30500]	1,18700	[0,74410 ; 1,7380]	0,14100	[0,03386 ; 0,24900]
35	0,94950	[0,50090 ; 1,66700]	0,47590	[-0,12730 ; 1,1880]	0,16820	[0,03189 ; 0,34160]
36	1,82500	[0,98140 ; 3,01800]	2,33000	[1,80400 ; 3,1000]	0,04620	[0,01182 ; 0,08515]

Fonte: Elaborado pela autora.

A Figura 4.2 apresenta as curvas características dos 34 itens considerados. Pode-se observar que, de maneira geral, elas possuem comportamento característico. Itens com valor do parâmetro de discriminação mais altos possuem curvas com maior inclinação e os itens com valor do parâmetro de discriminação mais baixos possuem inclinação mais suave, onde indivíduos com proficiência mais baixas e indivíduos com proficiência mais altas possuem praticamente a mesma probabilidade de resposta correta ao item. As curvas características posicionadas mais à direita do gráfico correspondem às curvas dos itens com parâmetros de dificuldade mais altos, onde a probabilidade de resposta correta é mais alta para indivíduos que possuem valor de proficiência mais alto. As curvas características mais à esquerda do gráfico correspondem aos itens com valor do parâmetro de dificuldade mais baixos, onde a probabilidade de resposta correta é alta para indivíduos com proficiência mais baixas. As curvas característica e de informação de cada item, assim como o percentual de respostas dadas para cada alternativa (correta e incorretas) pode ser consultado no Apêndice D.

Figura 4.2: CCI's para os 34 itens considerados.



Fonte: Elaborado pela autora.

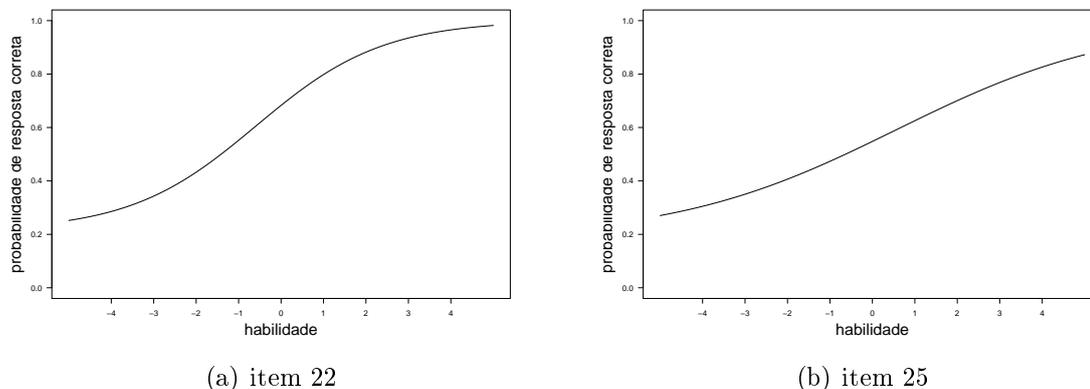
As Figuras 4.3 e 4.4 apresentam as curvas características dos itens 22 e 25 e 3 e 36 respectivamente. Pode-se observar nestas figuras as características citadas no parágrafo anterior. A curva Característica do item 22 por exemplo, apresentada na Figura 4.3 (a) que possui as estimativas $\hat{a}_{22} = 0,6713$, $\hat{b}_{22} = -0,5828$ e $\hat{c}_{22} = 0,2135$ permite observar que este pode ser considerado fácil, pois a probabilidade de resposta correta para indivíduos com proficiência -1 na escala $(0, 1)$ por exemplo, que corresponde a proficiência 50 na escala $(250, 50)$, é $0,55$.

Na Figura 4.3 (b), pode-se observar a curva característica do item 25, que possui as estimativas $\hat{a}_{25} = 0,3803$, $\hat{b}_{25} = 0,5593$ e $\hat{c}_{25} = 0,1823$. Este item possui baixo poder de discriminação. Pode-se observar que indivíduos com proficiências -1 e 1 na escala $(0, 1)$, que correspondem às proficiências 200 e 300 na escala $(250, 50)$, respectivamente, possuem probabilidade de resposta correta próximas $(0,47$ e $0,63)$.

Na Figura 4.4 (a) pode-se observar a curva característica do item 3 que possui estimativa para o parâmetro de dificuldade $\hat{b}_3 = -0,0085$. Para indivíduos com proficiência 1 na escala $(0, 1)$, correspondente a 300 na escala $(250, 50)$, a probabilidade de resposta correta a este item é $0,82$. Já em (b), onde o item 36 possui estimativa para o parâmetro

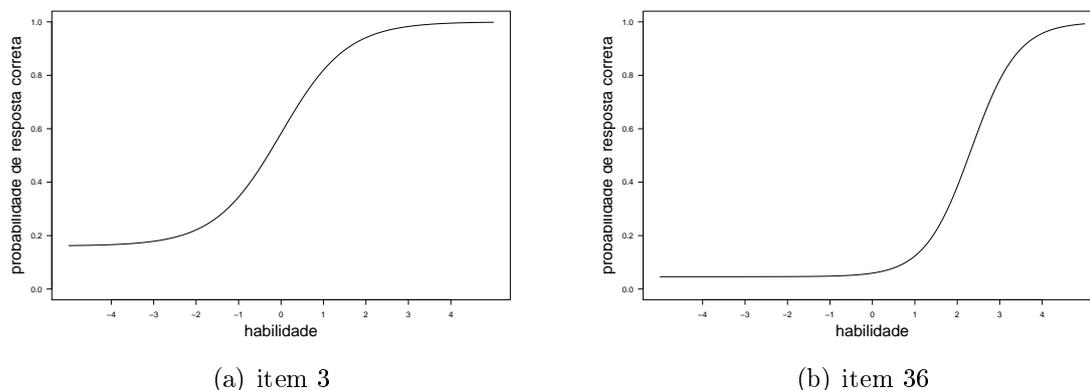
de dificuldade $\hat{b}_{36} = 2,33$, para a mesma probabilidade de resposta correta é necessário que a proficiência do indivíduo esteja próximo de 3, na escala (0, 1), o correspondente a 400 na escala (250, 50).

Figura 4.3: Curvas características dos itens 1 e 25.



Fonte: Elaborado pela autora.

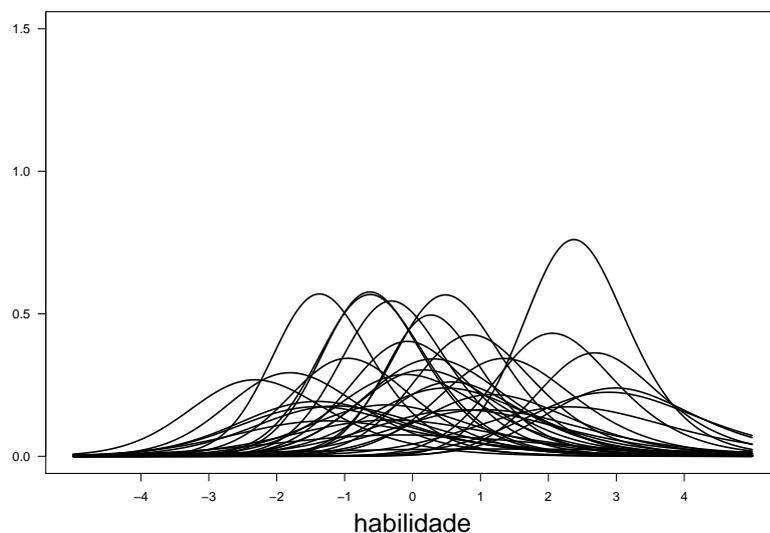
Figura 4.4: Curvas características dos itens 3 e 36.



Fonte: Elaborado pela autora.

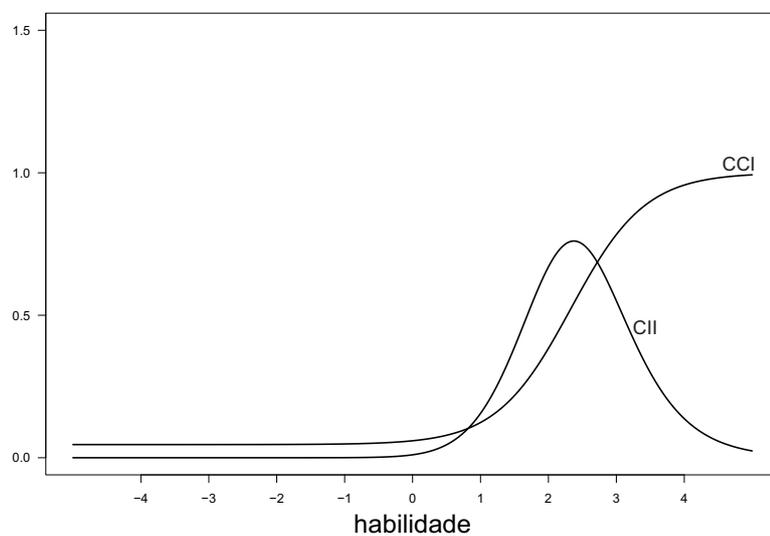
A Figura 4.5 apresenta as curvas de informação dos 34 itens considerados. A curva de informação do item permite analisar quanto um item contém de informação para a medida de proficiência. Como citado na Seção 3.2.2, para o modelo logístico de três parâmetros, a informação do item será maior quando b_i se aproxima de θ , quanto maior for o a_i e quanto mais c_i se aproximar de 0. A Figura 4.6 destaca a curva de informação e a curva característica do item 36 que possui estimativas dos parâmetros $\hat{a}_{36} = 1,825$, $\hat{b}_{36} = 2,33$ e $\hat{c}_{36} = 0,0462$. Este item possui o maior valor \hat{a} e o menor valor de \hat{c} entre os 34 itens considerados. Este é o que mais contém informação para a medida de proficiências em torno de 2 na escala (0, 1), que corresponde a 350 na escala (250, 50).

Figura 4.5: Curvas de informação dos itens.



Fonte: Elaborado pela autora.

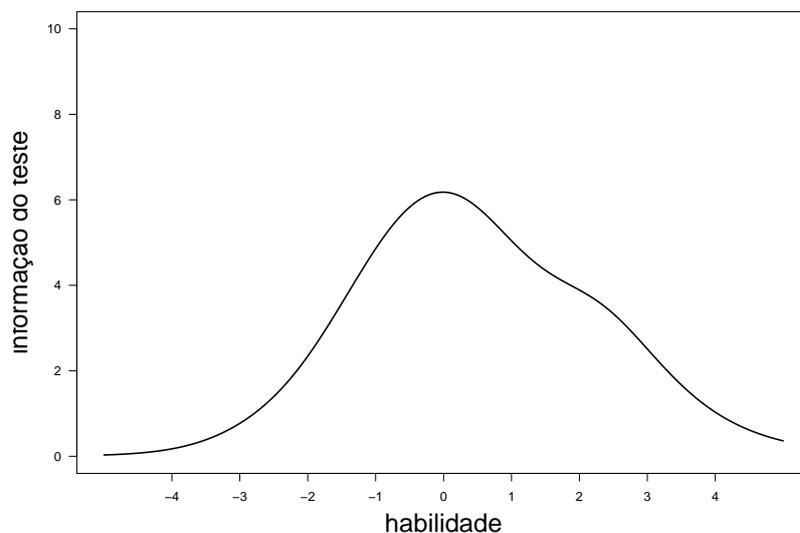
Figura 4.6: Curva de informação e curva característica do item 36.



Fonte: Elaborado pela autora.

A Figura 4.7 apresenta a Curva de Informação do Teste. A informação fornecida pelo teste é obtida pela soma das informações fornecidas pelos itens que o compõem e tem a característica de expressar o quão bem o teste está estimando as proficiências em toda a sua gama de pontuação. Pode-se observar nesta figura que a prova aplicada está melhor estimando as proficiências entre -2 e 3 na escala $(0, 1)$, que corresponde a 150 e 400 na escala $(250, 50)$, respectivamente.

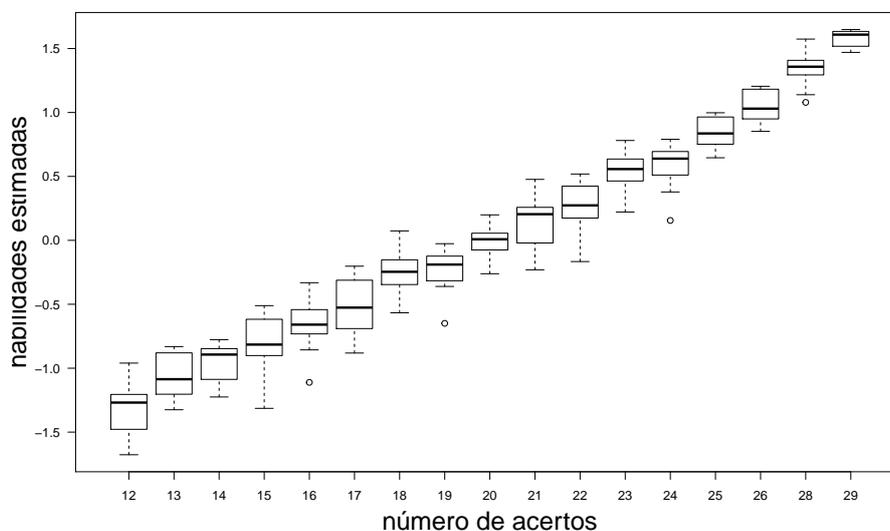
Figura 4.7: Curva de informação do teste.



Fonte: Elaborado pela autora.

A Tabela C.1 do Apêndice C apresenta as proficiências estimadas dos 277 respondentes. A tabela está em ordem crescente de acordo com a quantidade de acertos. Pode-se verificar que indivíduos com o mesmo número de acertos não possuem o mesmo valor de proficiência, pois como citado anteriormente, o cálculo da proficiência com base na TRI leva em consideração outros aspectos tais como, os parâmetros dos itens e o padrão de resposta dos indivíduos e não apenas a quantidade de acertos. Assim, a proficiência do indivíduo depende do seu padrão de resposta. Este fato também pode ser observado na Figura 4.8 que relaciona algumas das quantidades de número de acertos com as proficiências estimadas representadas por *boxplots*.

Figura 4.8: Relação entre números de acertos e proficiências estimadas.



Fonte: Elaborado pela autora.

4.4.3 Posicionamentos de Itens na Escala

Para tornar possível a interpretação pedagógica dos valores das proficiências, criou-se as escalas de proficiência também chamadas de escalas de conhecimento. As escalas de conhecimento são definidas a partir de níveis, conhecido como níveis âncora. Os níveis âncora por sua vez, são definidos por itens conhecidos como itens âncora. Estes níveis são pontos na escala, selecionados pelo pesquisador (analista), para serem interpretados pedagogicamente.

Como citado anteriormente, a escala de proficiência (250, 50) foi a adotada neste trabalho. A Tabela 4.4 apresenta a probabilidade de resposta correta para os 34 itens considerados nos níveis selecionados nesta escala e estão em um intervalo de 50 a 450, intercalados por 1 desvio padrão.

Após a escolha dos níveis âncora, para posicionar os itens âncora, utilizou-se inicialmente a proposta sugerida por Beaton e Allen (1992) apresentada na Seção 2.2. Pode-se observar que as três condições para posicionamento dos itens âncora só foram satisfeitas para 7 dos 34 considerados na prova. São os itens 2, 14, 15, 19, 20, 28, e 36 distribuídos entre os níveis 200 e 400 da escala (250, 50). Os itens e as probabilidades nos níveis em que foram posicionados como âncora, estão destacados em negrito (Tabela 4.4). O item 19 sugerido para avaliar a habilidade descrita na competência C19 da matriz de referência (Anexo 2), foi posicionado no nível 200. Os itens 14, 15 e 20, sugeridos para avaliar as habilidades descritas nas competências C14, C15 e C20, foram posicionados no nível 250. Os itens 2 e 28, sugeridos para avaliar as habilidades descritas nas competências C2 e C28 foram posicionados no nível 300. O item 36, sugerido para avaliar a habilidade descrita na competência C36, foi posicionado no nível 400.

Como a avaliação é composta de 36 itens, um para cada competência presente na matriz de referência, sendo considerados somente 34, ao utilizar o procedimento sugerido por Beaton e Allen (1992) não se consegue o posicionamento de itens âncora em todos os níveis escolhidos para interpretação da escala. Além disso, as habilidades cobertas pelos itens que não foram selecionados como âncora não estariam presentes na descrição dos níveis da escala. Dessa forma, através deste procedimento não é possível obter uma interpretação satisfatória para a escala de proficiência.

Tabela 4.4: Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por um desvio padrão.

Item	Níveis na escala								
	50	100	150	200	250	300	350	400	450
2	0,14	0,15	0,16	0,20	0,36	0,66	0,89	0,97	0,99
3	0,17	0,18	0,22	0,34	0,58	0,82	0,94	0,98	1,00
4	0,27	0,27	0,28	0,35	0,60	0,88	0,98	1,00	1,00
5	0,22	0,22	0,22	0,23	0,24	0,30	0,44	0,67	0,86
6	0,16	0,17	0,21	0,31	0,50	0,74	0,89	0,96	0,99
7	0,18	0,20	0,23	0,30	0,44	0,63	0,80	0,91	0,96
8	0,21	0,21	0,21	0,21	0,23	0,28	0,46	0,74	0,92
9	0,11	0,12	0,15	0,26	0,50	0,77	0,92	0,98	0,99
10	0,34	0,51	0,75	0,91	0,97	0,99	1,00	1,00	1,00
11	0,15	0,17	0,24	0,39	0,64	0,85	0,95	0,99	1,00
12	0,27	0,39	0,62	0,84	0,95	0,99	1,00	1,00	1,00
13	0,30	0,39	0,56	0,75	0,89	0,96	0,98	0,99	1,00
14	0,16	0,17	0,24	0,49	0,82	0,96	0,99	1,00	1,00
15	0,14	0,15	0,20	0,38	0,72	0,93	0,99	1,00	1,00
16	0,23	0,27	0,40	0,65	0,88	0,97	0,99	1,00	1,00
17	0,14	0,14	0,14	0,15	0,19	0,31	0,60	0,86	0,96
18	0,28	0,35	0,50	0,69	0,86	0,94	0,98	0,99	1,00
19	0,19	0,24	0,42	0,77	0,95	0,99	1,00	1,00	1,00
20	0,13	0,14	0,21	0,46	0,80	0,96	0,99	1,00	1,00
21	0,34	0,42	0,56	0,73	0,86	0,94	0,97	0,99	1,00
22	0,29	0,34	0,43	0,55	0,68	0,80	0,88	0,93	0,97
23	0,27	0,37	0,54	0,74	0,89	0,96	0,99	0,99	1,00
24	0,16	0,18	0,22	0,36	0,65	0,88	0,97	0,99	1,00
25	0,31	0,35	0,41	0,47	0,55	0,63	0,70	0,77	0,83
26	0,24	0,29	0,39	0,54	0,71	0,84	0,93	0,97	0,99
27	0,14	0,14	0,15	0,18	0,24	0,36	0,55	0,74	0,88
28	0,11	0,11	0,13	0,19	0,42	0,77	0,94	0,99	1,00
29	0,12	0,13	0,13	0,14	0,17	0,24	0,40	0,63	0,82
31	0,26	0,27	0,29	0,37	0,55	0,78	0,93	0,98	0,99
32	0,24	0,25	0,27	0,31	0,44	0,64	0,83	0,94	0,98
33	0,28	0,31	0,40	0,56	0,76	0,90	0,96	0,99	1,00
34	0,14	0,14	0,15	0,18	0,29	0,52	0,78	0,93	0,98
35	0,18	0,20	0,24	0,33	0,49	0,69	0,84	0,93	0,97
36	0,05	0,05	0,05	0,05	0,06	0,12	0,38	0,78	0,96

Fonte: Elaborado pela autora.

Uma segunda tentativa para o posicionamento dos itens âncora baseou-se na metodologia adotada pelo SAEB, apresentada na Subseção 2.2.1. Foram escolhidos níveis âncora na escala (250, 50), intercalados por meio desvio padrão, ou seja, com espaçamento 25, conforme Tabela 4.5, que aqui, apresenta a probabilidade de resposta correta calculada para os 34 itens considerados na prova, nos diferentes níveis selecionados. Neste caso, os níveis selecionados na escala (250, 50) estão em um intervalo de 75 a 425. Os itens foram posicionados como âncora no nível onde a probabilidade de resposta correta foi $\geq 0,65$ e se foi respondido corretamente por 50 ou mais respondentes. Pode-se observar que somente os itens 25 e 36 não foram posicionados como âncora. O item 25 não foi considerado bom

por possuir um valor de parâmetro a baixo e o item 36 foi respondido corretamente por apenas 27 indivíduos indicando que este é um item muito difícil. Conforme mencionado anteriormente, tanto os itens quando as probabilidades de resposta correta nos níveis onde foram posicionados estão destacados em negrito.

Tabela 4.5: Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por meio desvio padrão.

Item	Níveis na escala intercalados por meio desvio padrão														
	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425
2	0,14	0,15	0,15	0,16	0,17	0,20	0,26	0,36	0,50	0,66	0,79	0,89	0,94	0,97	0,99
3	0,17	0,18	0,19	0,22	0,27	0,34	0,45	0,58	0,71	0,82	0,89	0,94	0,97	0,98	0,99
4	0,27	0,27	0,27	0,28	0,30	0,35	0,45	0,60	0,76	0,88	0,95	0,98	0,99	1,00	1,00
5	0,22	0,22	0,22	0,22	0,22	0,23	0,23	0,24	0,27	0,30	0,36	0,44	0,55	0,67	0,77
6	0,16	0,17	0,19	0,21	0,25	0,31	0,40	0,50	0,62	0,74	0,83	0,89	0,94	0,96	0,98
7	0,19	0,20	0,21	0,23	0,26	0,30	0,36	0,44	0,53	0,63	0,72	0,80	0,86	0,91	0,94
8	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,23	0,25	0,28	0,35	0,46	0,60	0,74	0,85
9	0,11	0,12	0,13	0,15	0,19	0,26	0,37	0,50	0,64	0,77	0,86	0,92	0,96	0,98	0,99
10	0,41	0,51	0,63	0,75	0,84	0,91	0,95	0,97	0,99	0,99	1,00	1,00	1,00	1,00	1,00
11	0,16	0,17	0,20	0,24	0,30	0,39	0,51	0,64	0,76	0,85	0,91	0,95	0,97	0,99	0,99
12	0,32	0,39	0,49	0,62	0,74	0,84	0,91	0,95	0,97	0,99	0,99	1,00	1,00	1,00	1,00
13	0,34	0,39	0,47	0,56	0,66	0,75	0,83	0,89	0,93	0,96	0,97	0,98	0,99	0,99	1,00
14	0,16	0,17	0,19	0,24	0,34	0,49	0,67	0,82	0,92	0,96	0,98	0,99	1,00	1,00	1,00
15	0,15	0,15	0,17	0,20	0,26	0,38	0,55	0,72	0,85	0,93	0,97	0,99	0,99	1,00	1,00
16	0,24	0,27	0,32	0,40	0,51	0,65	0,78	0,88	0,93	0,97	0,98	0,99	1,00	1,00	1,00
17	0,14	0,14	0,14	0,14	0,15	0,15	0,16	0,19	0,23	0,31	0,44	0,60	0,75	0,86	0,93
18	0,31	0,35	0,41	0,50	0,59	0,69	0,78	0,86	0,91	0,94	0,96	0,98	0,99	0,99	1,00
19	0,21	0,24	0,30	0,42	0,60	0,77	0,89	0,95	0,98	0,99	1,00	1,00	1,00	1,00	1,00
20	0,13	0,14	0,16	0,21	0,31	0,46	0,65	0,80	0,90	0,96	0,98	0,99	1,00	1,00	1,00
21	0,37	0,42	0,49	0,56	0,65	0,73	0,80	0,86	0,90	0,94	0,96	0,97	0,98	0,99	0,99
22	0,31	0,34	0,38	0,43	0,49	0,55	0,62	0,68	0,74	0,80	0,84	0,88	0,91	0,93	0,95
23	0,31	0,37	0,44	0,54	0,64	0,74	0,83	0,89	0,93	0,96	0,98	0,99	0,99	0,99	1,00
24	0,17	0,18	0,19	0,22	0,27	0,36	0,50	0,65	0,79	0,88	0,94	0,97	0,99	0,99	1,00
25	0,33	0,35	0,38	0,41	0,44	0,47	0,51	0,55	0,59	0,63	0,66	0,70	0,74	0,77	0,80
26	0,26	0,29	0,34	0,39	0,46	0,54	0,62	0,71	0,78	0,84	0,89	0,93	0,95	0,97	0,98
27	0,14	0,14	0,14	0,15	0,16	0,18	0,20	0,24	0,29	0,36	0,45	0,55	0,65	0,74	0,82
28	0,11	0,11	0,12	0,13	0,15	0,19	0,28	0,42	0,60	0,77	0,88	0,94	0,98	0,99	1,00
29	0,12	0,13	0,13	0,13	0,13	0,14	0,15	0,17	0,20	0,24	0,31	0,40	0,51	0,63	0,74
31	0,27	0,27	0,28	0,29	0,32	0,37	0,45	0,55	0,67	0,78	0,87	0,93	0,96	0,98	0,99
32	0,25	0,25	0,26	0,27	0,28	0,31	0,36	0,44	0,53	0,64	0,75	0,83	0,90	0,94	0,97
33	0,29	0,31	0,35	0,40	0,47	0,56	0,66	0,76	0,84	0,90	0,94	0,96	0,98	0,99	0,99
34	0,14	0,14	0,15	0,15	0,16	0,18	0,22	0,29	0,39	0,52	0,66	0,78	0,87	0,93	0,96
35	0,19	0,20	0,21	0,24	0,28	0,33	0,40	0,49	0,59	0,69	0,77	0,84	0,89	0,93	0,96
36	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,06	0,08	0,12	0,22	0,38	0,60	0,78	0,90

Fonte: Elaborado pela autora.

Uma terceira tentativa para o posicionamento dos itens âncora baseou-se na metodologia adotada pelo SARESP, apresentada na Subseção 2.2.3. Os itens foram posicionados como âncora se a probabilidade de resposta correta no nível considerado foi $\geq 0,65$ e a probabilidade de resposta correta dos itens no nível dois espaçamentos anteriores ao nível considerado foi $< 0,50$. Foram escolhidos níveis âncora da mesma forma como foi feito na Tabela 4.5. Os itens e as probabilidades nos níveis em que foram posicionados como âncora também estão destacados em negrito.

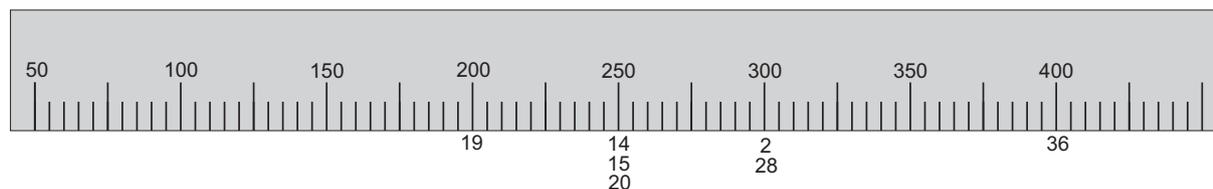
Tabela 4.6: Probabilidade de resposta correta para cada um dos itens considerados, condicionada ao nível de proficiência na escala intercalados por meio desvio padrão.

Item	Níveis na escala intercalados por meio desvio padrão														
	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425
2	0,14	0,15	0,15	0,16	0,17	0,20	0,26	0,36	0,50	0,66	0,79	0,89	0,94	0,97	0,99
3	0,17	0,18	0,19	0,22	0,27	0,34	0,45	0,58	0,71	0,82	0,89	0,94	0,97	0,98	0,99
4	0,27	0,27	0,27	0,28	0,30	0,35	0,45	0,60	0,76	0,88	0,95	0,98	0,99	1,00	1,00
5	0,22	0,22	0,22	0,22	0,22	0,23	0,23	0,24	0,27	0,30	0,36	0,44	0,55	0,67	0,77
6	0,16	0,17	0,19	0,21	0,25	0,31	0,40	0,50	0,62	0,74	0,83	0,89	0,94	0,96	0,98
7	0,19	0,20	0,21	0,23	0,26	0,30	0,36	0,44	0,53	0,63	0,72	0,80	0,86	0,91	0,94
8	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,23	0,25	0,28	0,35	0,46	0,60	0,74	0,85
9	0,11	0,12	0,13	0,15	0,19	0,26	0,37	0,50	0,64	0,77	0,86	0,92	0,96	0,98	0,99
10	0,41	0,51	0,63	0,75	0,84	0,91	0,95	0,97	0,99	0,99	1,00	1,00	1,00	1,00	1,00
11	0,16	0,17	0,20	0,24	0,30	0,39	0,51	0,64	0,76	0,85	0,91	0,95	0,97	0,99	0,99
12	0,32	0,39	0,49	0,62	0,74	0,84	0,91	0,95	0,97	0,99	0,99	1,00	1,00	1,00	1,00
13	0,34	0,39	0,47	0,56	0,66	0,75	0,83	0,89	0,93	0,96	0,97	0,98	0,99	0,99	1,00
14	0,16	0,17	0,19	0,24	0,34	0,49	0,67	0,82	0,92	0,96	0,98	0,99	1,00	1,00	1,00
15	0,15	0,15	0,17	0,20	0,26	0,38	0,55	0,72	0,85	0,93	0,97	0,99	0,99	1,00	1,00
16	0,24	0,27	0,32	0,40	0,51	0,65	0,78	0,88	0,93	0,97	0,98	0,99	1,00	1,00	1,00
17	0,14	0,14	0,14	0,14	0,15	0,15	0,16	0,19	0,23	0,31	0,44	0,60	0,75	0,86	0,93
18	0,31	0,35	0,41	0,50	0,59	0,69	0,78	0,86	0,91	0,94	0,96	0,98	0,99	0,99	1,00
19	0,21	0,24	0,30	0,42	0,60	0,77	0,89	0,95	0,98	0,99	1,00	1,00	1,00	1,00	1,00
20	0,13	0,14	0,16	0,21	0,31	0,46	0,65	0,80	0,90	0,96	0,98	0,99	1,00	1,00	1,00
21	0,37	0,42	0,49	0,56	0,65	0,73	0,80	0,86	0,90	0,94	0,96	0,97	0,98	0,99	0,99
22	0,31	0,34	0,38	0,43	0,49	0,55	0,62	0,68	0,74	0,80	0,84	0,88	0,91	0,93	0,95
23	0,31	0,37	0,44	0,54	0,64	0,74	0,83	0,89	0,93	0,96	0,98	0,99	0,99	0,99	1,00
24	0,17	0,18	0,19	0,22	0,27	0,36	0,50	0,65	0,79	0,88	0,94	0,97	0,99	0,99	1,00
25	0,33	0,35	0,38	0,41	0,44	0,47	0,51	0,55	0,59	0,63	0,66	0,70	0,74	0,77	0,80
26	0,26	0,29	0,34	0,39	0,46	0,54	0,62	0,71	0,78	0,84	0,89	0,93	0,95	0,97	0,98
27	0,14	0,14	0,14	0,15	0,16	0,18	0,20	0,24	0,29	0,36	0,45	0,55	0,65	0,74	0,82
28	0,11	0,11	0,12	0,13	0,15	0,19	0,28	0,42	0,60	0,77	0,88	0,94	0,98	0,99	1,00
29	0,12	0,13	0,13	0,13	0,13	0,14	0,15	0,17	0,20	0,24	0,31	0,40	0,51	0,63	0,74
31	0,27	0,27	0,28	0,29	0,32	0,37	0,45	0,55	0,67	0,78	0,87	0,93	0,96	0,98	0,99
32	0,25	0,25	0,26	0,27	0,28	0,31	0,36	0,44	0,53	0,64	0,75	0,83	0,90	0,94	0,97
33	0,29	0,31	0,35	0,40	0,47	0,56	0,66	0,76	0,84	0,90	0,94	0,96	0,98	0,99	0,99
34	0,14	0,14	0,15	0,15	0,16	0,18	0,22	0,29	0,39	0,52	0,66	0,78	0,87	0,93	0,96
35	0,19	0,20	0,21	0,24	0,28	0,33	0,40	0,49	0,59	0,69	0,77	0,84	0,89	0,93	0,96
36	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,06	0,08	0,12	0,22	0,38	0,60	0,78	0,90

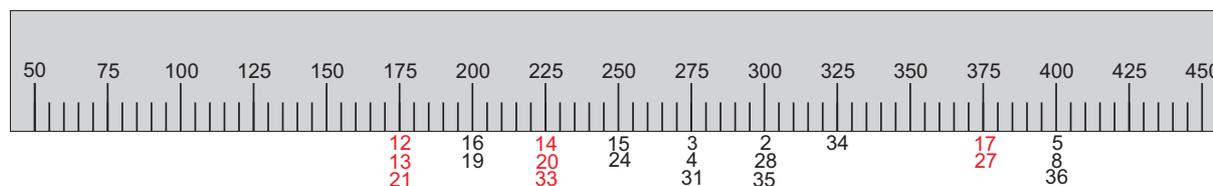
Fonte: Elaborado pela autora.

Comparando o posicionamento dos itens âncora, nos níveis âncora, na escala representada pelas régua na Figura 4.9, pode-se observar em (a), posicionamento sugerido por Beaton e Allen (1992), que poucos itens foram posicionados e em poucos níveis da escala, o que impossibilita uma boa interpretação da mesma. Observando a régua apresentada em (b), posicionamento baseado no SARESP, verifica-se que dos 34 itens considerados 22 foram posicionados. Ainda assim, em alguns níveis da escala não foram posicionados item, o que dificulta uma boa interpretação da escala. Observando a régua apresentada em (c), posicionamento baseado no SAEB, verifica-se que dos 34 itens 32 foram posicionados. Observa-se aqui também níveis sem o posicionamento de itens. Comparando (b) e (c) pode-se observar que em qualquer um destes dois procedimentos, a interpretação dos níveis 175, 225 e 375 seria a mesma, pois temos os mesmos itens posicionados destacados em vermelho.

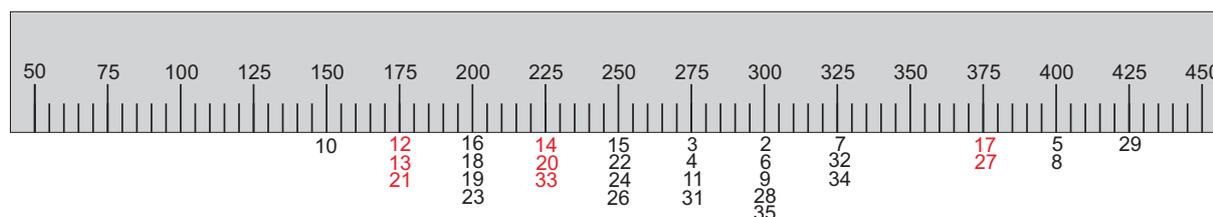
Figura 4.9: Posicionamento dos itens âncora nos níveis âncora da escala representados por régua, de acordo com o sugerido por Beaton e Allen (1992) e baseados nos procedimentos utilizado pelo SAEB e SARESP.



(a) Posicionamento sugerido por Beaton e Allen (1992)



(b) Posicionamento baseado no utilizado pelo SARESP



(c) Posicionamento baseado no utilizado pelo SAEB

Fonte: Elaborado pela autora.

Pelas análises apresentadas pode-se observar que nenhuma das três metodologias permitiu uma boa interpretação em toda a extensão da escala, pois existem níveis onde não foram posicionados itens âncora impossibilitando a interpretação do mesmo. Optou-se, então, pela técnica adotada por Barros (2016), que consiste em dividir os níveis da escala em grupos, a partir da análise de agrupamentos hierárquicos, segmentando a escala em faixas de proficiências, conforme descrito a seguir.

4.5 Construção e Interpretação da Escala de Proficiência em Conteúdos Matemáticos Básicos

A construção da escala envolve a segmentação e o posicionamento dos itens. A segmentação da escala é apresentada na Subseção 4.5.1. Na Subseção 4.5.2 é apresentada a interpretação da escala.

4.5.1 Construção da Escala

O procedimento aqui utilizado para segmentação da escala baseou-se na técnica adotada por Barros (2016). Para isso, definiu-se valores de proficiências no intervalo de 75 a 425 (na escala) com diferença de 5 unidades entre um e outro, totalizando 71 valores.

Calculou-se então, a probabilidade de resposta correta para estes valores para cada item conforme Tabelas 4.9, 4.10, 4.11, 4.12 e 4.13.

Após o cálculo das probabilidades, para cada item nos 71 níveis, realizou-se a análise de agrupamentos. Como variáveis de entrada utilizou-se os 71 valores de proficiência selecionados e as probabilidades de resposta correta para cada item, em cada um desses valores.

A análise de agrupamentos hierárquicos foi realizada no software IBM SPSS Statistics 23 a partir do método de Ward, onde as variáveis de entrada foram dispostas da seguinte forma: os itens em colunas e os valores das probabilidades em linhas. O agrupamento dos níveis pode ser observado no dendograma apresentado Figura 4.10.

A partir da análise do dendograma optou-se inicialmente pela escolha de 4 grupos. Selecionados os grupos, se faz necessário o posicionamento dos itens como âncora para que seja possível a interpretação da escala em cada grupo selecionado. Para o posicionamento do item como âncora nos grupos escolhidos foi utilizado o seguinte procedimento: o item foi posicionado no nível onde a probabilidade de resposta correta foi $\geq 60\%$ e o valor do parâmetro $a \geq 1$ para cada item. A Tabela 4.7 apresenta o número de grupos selecionados, o intervalo da escala para cada grupo, a porcentagem de alunos em cada grupo e o número de itens âncora posicionados.

Tabela 4.7: Número de grupos selecionados, intervalo da escala para cada grupo, porcentagem de alunos em cada grupo e o número de itens âncora posicionados.

Grupos	Intervalo da escala	% de alunos	Nº itens âncora posicionados
Grupo 1	[75 a 155]	1,08	2
Grupo 2]155 à 230]	31,77	8
Grupo 3]230 à 335]	64,26	13
Grupo 4]335 à 425]	2,89	5

Fonte: Elaborado pela autora.

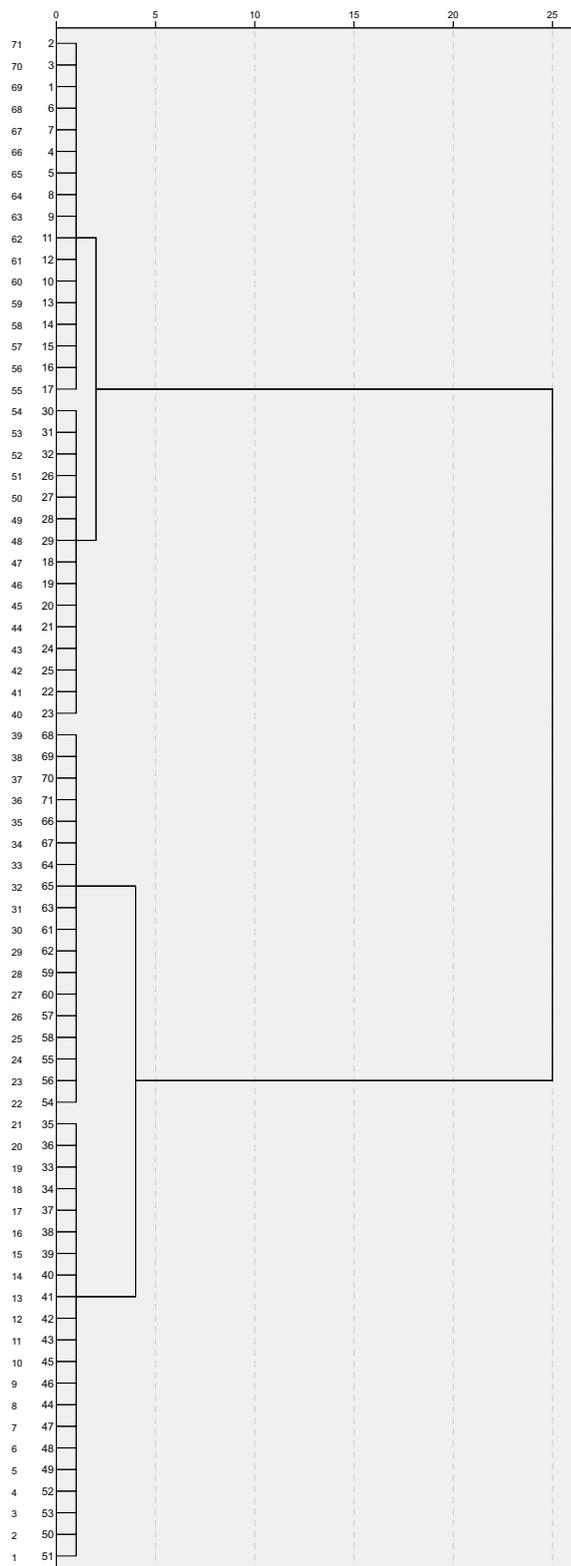
Pode-se observar a partir da Tabela 4.7 que o grupo 3 possui a maior quantidade de itens posicionados como âncora, por esta razão, buscando uma melhor interpretação da escala em toda a sua amplitude, optou-se por fazer o mesmo procedimento relatado, de agrupamentos hierárquicos, considerando apenas os valores das proficiências e os valores de probabilidade referentes ao grupo 3, buscando uma nova proposta de divisão para esse grupo. Analisando os resultados obtidos optou-se por dividir este grupo em outros dois e passou-se a ter a estrutura apresentada na Tabela 4.8.

Tabela 4.8: Número de grupos selecionados, intervalo da escala para cada grupo, porcentagem de alunos em cada grupo e o número de itens âncora posicionados.

Grupos	Intervalo da escala	% de alunos	Nº itens âncora posicionados
Grupo 1	[75 à 155]	1,08	2
Grupo 2]155 à 230]	31,77	8
Grupo 3]230 à 275]	34,66	9
Grupo 4]275 à 335]	29,60	4
Grupo 5]335 à 425]	2,89	5

Fonte: Elaborado pela autora.

Figura 4.10: Dendograma.



Fonte: Elaborado pela autora.

Nas Tabelas 4.9, 4.10, 4.11, 4.12 e 4.13 podem ser observados o posicionamento dos itens âncora nos níveis na escala. Os itens e as probabilidades nos níveis em que foram posicionados como âncora estão destacados em negrito.

Tabela 4.9: Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 1.

Item	Níveis na escala - Grupo 1																
	75	80	85	90	95	100	105	110	115	120	125	130	135	140	145	150	155
2	0,14	0,14	0,14	0,14	0,14	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,16	0,16
3	0,17	0,17	0,17	0,17	0,18	0,18	0,18	0,18	0,19	0,19	0,19	0,20	0,20	0,20	0,21	0,22	0,23
4	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,28	0,28	0,28	0,28	0,28
5	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22
6	0,16	0,16	0,17	0,17	0,17	0,17	0,18	0,18	0,18	0,18	0,19	0,19	0,19	0,20	0,20	0,21	0,22
7	0,19	0,19	0,19	0,19	0,19	0,20	0,20	0,20	0,20	0,21	0,21	0,21	0,21	0,22	0,22	0,23	0,23
8	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21
9	0,11	0,11	0,11	0,11	0,11	0,12	0,12	0,12	0,12	0,13	0,13	0,13	0,14	0,14	0,15	0,15	0,16
10	0,41	0,43	0,44	0,46	0,49	0,51	0,53	0,56	0,58	0,61	0,63	0,65	0,68	0,70	0,73	0,75	0,77
11	0,16	0,16	0,17	0,17	0,17	0,17	0,18	0,18	0,19	0,19	0,20	0,20	0,21	0,22	0,23	0,24	0,25
12	0,32	0,33	0,34	0,35	0,37	0,39	0,40	0,42	0,44	0,47	0,49	0,51	0,54	0,56	0,59	0,62	0,64
13	0,34	0,35	0,36	0,37	0,38	0,39	0,40	0,42	0,43	0,45	0,47	0,48	0,50	0,52	0,54	0,56	0,58
14	0,16	0,16	0,17	0,17	0,17	0,17	0,18	0,18	0,18	0,19	0,19	0,20	0,21	0,22	0,23	0,24	0,26
15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,16	0,16	0,16	0,17	0,17	0,18	0,18	0,19	0,20	0,21
16	0,24	0,25	0,25	0,26	0,26	0,27	0,28	0,28	0,29	0,30	0,32	0,33	0,34	0,36	0,38	0,40	0,42
17	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14
18	0,31	0,31	0,32	0,33	0,34	0,35	0,36	0,37	0,39	0,40	0,41	0,43	0,45	0,46	0,48	0,50	0,52
19	0,21	0,21	0,21	0,22	0,23	0,24	0,25	0,26	0,27	0,28	0,30	0,32	0,34	0,37	0,39	0,42	0,45
20	0,13	0,13	0,13	0,14	0,14	0,14	0,14	0,15	0,15	0,16	0,16	0,17	0,18	0,19	0,20	0,21	0,23
21	0,37	0,38	0,39	0,40	0,41	0,42	0,43	0,45	0,46	0,47	0,49	0,50	0,52	0,53	0,55	0,56	0,58
22	0,31	0,32	0,32	0,33	0,34	0,34	0,35	0,36	0,37	0,37	0,38	0,39	0,40	0,41	0,42	0,43	0,44
23	0,31	0,32	0,33	0,34	0,35	0,37	0,38	0,39	0,41	0,43	0,44	0,46	0,48	0,50	0,52	0,54	0,56
24	0,17	0,17	0,17	0,17	0,17	0,18	0,18	0,18	0,18	0,19	0,19	0,19	0,20	0,20	0,21	0,22	0,23
25	0,33	0,33	0,34	0,34	0,35	0,35	0,36	0,36	0,37	0,37	0,38	0,38	0,39	0,39	0,40	0,41	0,41
26	0,26	0,27	0,28	0,28	0,29	0,29	0,30	0,31	0,32	0,33	0,34	0,34	0,36	0,37	0,38	0,39	0,40
27	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,15	0,15	0,15	0,15	0,15
28	0,11	0,11	0,11	0,11	0,11	0,11	0,12	0,12	0,12	0,12	0,12	0,12	0,12	0,12	0,13	0,13	0,13
29	0,12	0,12	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13	0,13
31	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,27	0,28	0,28	0,28	0,28	0,29	0,29	0,29	0,30
32	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,26	0,26	0,26	0,26	0,26	0,27	0,27
33	0,29	0,30	0,30	0,30	0,31	0,31	0,32	0,33	0,33	0,34	0,35	0,36	0,37	0,38	0,39	0,40	0,41
34	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15
35	0,19	0,19	0,19	0,19	0,20	0,20	0,20	0,21	0,21	0,21	0,21	0,22	0,22	0,23	0,23	0,24	0,25
36	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05

Fonte: Elaborado pela autora.

Tabela 4.10: Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 2.

Item	Níveis na escala - Grupo 2														
	160	165	170	175	180	185	190	195	200	205	210	215	220	225	230
2	0,16	0,16	0,17	0,17	0,18	0,18	0,19	0,19	0,20	0,21	0,22	0,23	0,25	0,26	0,28
3	0,24	0,25	0,26	0,27	0,28	0,30	0,31	0,33	0,34	0,36	0,38	0,41	0,43	0,45	0,48
4	0,29	0,29	0,30	0,30	0,31	0,32	0,33	0,34	0,35	0,37	0,39	0,41	0,43	0,45	0,48
5	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,22	0,23	0,23	0,23	0,23	0,23	0,23	0,23
6	0,22	0,23	0,24	0,25	0,26	0,27	0,28	0,30	0,31	0,32	0,34	0,36	0,38	0,40	0,42
7	0,24	0,24	0,25	0,26	0,27	0,27	0,28	0,29	0,30	0,31	0,32	0,33	0,35	0,36	0,37
8	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,21	0,22
9	0,17	0,17	0,18	0,19	0,21	0,22	0,23	0,25	0,26	0,28	0,30	0,32	0,34	0,37	0,39
10	0,79	0,81	0,83	0,84	0,86	0,87	0,89	0,90	0,91	0,92	0,93	0,94	0,94	0,95	0,96
11	0,26	0,27	0,28	0,30	0,32	0,33	0,35	0,37	0,39	0,41	0,44	0,46	0,49	0,51	0,54
12	0,67	0,69	0,72	0,74	0,76	0,79	0,81	0,82	0,84	0,86	0,87	0,89	0,90	0,91	0,92
13	0,60	0,62	0,64	0,66	0,68	0,69	0,71	0,73	0,75	0,77	0,78	0,80	0,82	0,83	0,84
14	0,27	0,29	0,31	0,34	0,36	0,39	0,42	0,45	0,49	0,52	0,56	0,60	0,64	0,67	0,71
15	0,22	0,23	0,25	0,26	0,28	0,30	0,33	0,35	0,38	0,41	0,44	0,48	0,51	0,55	0,58
16	0,44	0,46	0,49	0,51	0,54	0,57	0,60	0,62	0,65	0,68	0,71	0,73	0,76	0,78	0,80
17	0,14	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,16	0,16	0,16	0,16	0,17
18	0,53	0,55	0,57	0,59	0,61	0,63	0,65	0,67	0,69	0,71	0,73	0,75	0,77	0,78	0,80
19	0,49	0,52	0,56	0,60	0,63	0,67	0,70	0,74	0,77	0,80	0,82	0,85	0,87	0,89	0,90
20	0,25	0,26	0,29	0,31	0,34	0,37	0,40	0,43	0,46	0,50	0,54	0,57	0,61	0,65	0,68
21	0,60	0,61	0,63	0,65	0,66	0,68	0,70	0,71	0,73	0,74	0,76	0,77	0,79	0,80	0,81
22	0,45	0,47	0,48	0,49	0,50	0,51	0,53	0,54	0,55	0,57	0,58	0,59	0,60	0,62	0,63
23	0,58	0,60	0,62	0,64	0,66	0,68	0,70	0,72	0,74	0,76	0,78	0,80	0,81	0,83	0,84
24	0,24	0,25	0,26	0,27	0,29	0,30	0,32	0,34	0,36	0,39	0,41	0,44	0,47	0,50	0,53
25	0,42	0,43	0,43	0,44	0,45	0,45	0,46	0,47	0,47	0,48	0,49	0,50	0,50	0,51	0,52
26	0,42	0,43	0,44	0,46	0,47	0,49	0,51	0,52	0,54	0,56	0,57	0,59	0,61	0,62	0,64
27	0,15	0,16	0,16	0,16	0,16	0,17	0,17	0,17	0,18	0,18	0,19	0,19	0,20	0,20	0,21
28	0,13	0,14	0,14	0,15	0,15	0,16	0,17	0,18	0,19	0,20	0,22	0,24	0,26	0,28	0,30
29	0,13	0,13	0,13	0,13	0,13	0,14	0,14	0,14	0,14	0,14	0,14	0,15	0,15	0,15	0,15
31	0,30	0,31	0,32	0,32	0,33	0,34	0,35	0,36	0,37	0,38	0,40	0,41	0,43	0,45	0,47
32	0,27	0,28	0,28	0,28	0,29	0,29	0,30	0,31	0,31	0,32	0,33	0,34	0,35	0,36	0,38
33	0,43	0,44	0,46	0,47	0,49	0,51	0,53	0,55	0,56	0,58	0,60	0,62	0,64	0,66	0,68
34	0,16	0,16	0,16	0,16	0,17	0,17	0,17	0,18	0,18	0,19	0,20	0,20	0,21	0,22	0,23
35	0,25	0,26	0,27	0,28	0,29	0,30	0,31	0,32	0,33	0,35	0,36	0,37	0,39	0,40	0,42
36	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05

Fonte: Elaborado pela autora.

Tabela 4.11: Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 3.

Item	Níveis na escala - Grupo 3								
	235	240	245	250	255	260	265	270	275
2	0,29	0,31	0,33	0,36	0,38	0,41	0,44	0,47	0,50
3	0,50	0,53	0,56	0,58	0,61	0,64	0,66	0,69	0,71
4	0,51	0,54	0,57	0,60	0,64	0,67	0,70	0,73	0,76
5	0,24	0,24	0,24	0,24	0,25	0,25	0,26	0,26	0,27
6	0,44	0,46	0,48	0,50	0,53	0,55	0,58	0,60	0,62
7	0,39	0,40	0,42	0,44	0,45	0,47	0,49	0,51	0,53
8	0,22	0,22	0,22	0,23	0,23	0,23	0,24	0,24	0,25
9	0,42	0,44	0,47	0,50	0,53	0,56	0,59	0,61	0,64
10	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,99
11	0,56	0,59	0,62	0,64	0,67	0,69	0,72	0,74	0,76
12	0,93	0,94	0,94	0,95	0,96	0,96	0,97	0,97	0,97
13	0,86	0,87	0,88	0,89	0,90	0,91	0,92	0,92	0,93
14	0,74	0,77	0,80	0,82	0,85	0,87	0,89	0,90	0,92
15	0,62	0,66	0,69	0,72	0,75	0,78	0,81	0,83	0,85
16	0,82	0,84	0,86	0,88	0,89	0,90	0,91	0,93	0,93
17	0,17	0,18	0,18	0,19	0,19	0,20	0,21	0,22	0,23
18	0,81	0,83	0,84	0,86	0,87	0,88	0,89	0,90	0,91
19	0,92	0,93	0,94	0,95	0,96	0,96	0,97	0,97	0,98
20	0,72	0,75	0,78	0,80	0,83	0,85	0,87	0,89	0,90
21	0,83	0,84	0,85	0,86	0,87	0,88	0,89	0,90	0,90
22	0,64	0,66	0,67	0,68	0,70	0,71	0,72	0,73	0,74
23	0,85	0,87	0,88	0,89	0,90	0,91	0,92	0,92	0,93
24	0,56	0,59	0,62	0,65	0,68	0,71	0,74	0,76	0,79
25	0,52	0,53	0,54	0,55	0,56	0,56	0,57	0,58	0,59
26	0,66	0,67	0,69	0,71	0,72	0,74	0,75	0,77	0,78
27	0,21	0,22	0,23	0,24	0,25	0,26	0,27	0,28	0,29
28	0,33	0,35	0,39	0,42	0,45	0,49	0,53	0,56	0,60
29	0,16	0,16	0,16	0,17	0,17	0,18	0,19	0,19	0,20
31	0,49	0,51	0,53	0,55	0,58	0,60	0,62	0,65	0,67
32	0,39	0,40	0,42	0,44	0,45	0,47	0,49	0,51	0,53
33	0,70	0,72	0,74	0,76	0,78	0,79	0,81	0,82	0,84
34	0,24	0,26	0,27	0,29	0,30	0,32	0,34	0,36	0,39
35	0,44	0,46	0,47	0,49	0,51	0,53	0,55	0,57	0,59
36	0,05	0,06	0,06	0,06	0,06	0,07	0,07	0,07	0,08

Fonte: Elaborado pela autora.

Tabela 4.12: Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 4.

Item	Níveis na escala - Grupo 4											
	280	285	290	295	300	305	310	315	320	325	330	335
2	0,53	0,56	0,59	0,63	0,66	0,69	0,72	0,74	0,77	0,79	0,82	0,84
3	0,74	0,76	0,78	0,80	0,82	0,84	0,85	0,87	0,88	0,89	0,91	0,92
4	0,79	0,82	0,84	0,86	0,88	0,90	0,91	0,93	0,94	0,95	0,96	0,96
5	0,27	0,28	0,29	0,29	0,30	0,31	0,32	0,33	0,35	0,36	0,37	0,39
6	0,65	0,67	0,69	0,71	0,74	0,76	0,77	0,79	0,81	0,83	0,84	0,86
7	0,55	0,57	0,59	0,61	0,63	0,65	0,67	0,68	0,70	0,72	0,74	0,76
8	0,25	0,26	0,27	0,27	0,28	0,30	0,31	0,32	0,34	0,35	0,37	0,39
9	0,67	0,70	0,72	0,75	0,77	0,79	0,81	0,83	0,85	0,86	0,88	0,89
10	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00
11	0,78	0,80	0,82	0,84	0,85	0,87	0,88	0,89	0,90	0,91	0,92	0,93
12	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
13	0,94	0,94	0,95	0,95	0,96	0,96	0,96	0,97	0,97	0,97	0,98	0,98
14	0,93	0,94	0,95	0,96	0,96	0,97	0,97	0,98	0,98	0,98	0,99	0,99
15	0,87	0,89	0,91	0,92	0,93	0,94	0,95	0,96	0,96	0,97	0,97	0,98
16	0,94	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,99	0,99
17	0,25	0,26	0,28	0,29	0,31	0,34	0,36	0,38	0,41	0,44	0,47	0,50
18	0,92	0,92	0,93	0,94	0,94	0,95	0,95	0,96	0,96	0,96	0,97	0,97
19	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00
20	0,92	0,93	0,94	0,95	0,96	0,96	0,97	0,97	0,98	0,98	0,98	0,99
21	0,91	0,92	0,92	0,93	0,94	0,94	0,95	0,95	0,95	0,96	0,96	0,96
22	0,76	0,77	0,78	0,79	0,80	0,81	0,82	0,83	0,84	0,84	0,85	0,86
23	0,94	0,94	0,95	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98
24	0,81	0,83	0,85	0,87	0,88	0,90	0,91	0,92	0,93	0,94	0,95	0,95
25	0,59	0,60	0,61	0,62	0,63	0,63	0,64	0,65	0,66	0,66	0,67	0,68
26	0,80	0,81	0,82	0,83	0,84	0,85	0,86	0,87	0,88	0,89	0,90	0,91
27	0,30	0,32	0,33	0,35	0,36	0,38	0,39	0,41	0,43	0,45	0,47	0,49
28	0,64	0,67	0,71	0,74	0,77	0,80	0,82	0,84	0,86	0,88	0,90	0,91
29	0,21	0,21	0,22	0,23	0,24	0,26	0,27	0,28	0,30	0,31	0,33	0,34
31	0,70	0,72	0,74	0,76	0,78	0,80	0,82	0,84	0,85	0,87	0,88	0,90
32	0,55	0,57	0,60	0,62	0,64	0,66	0,68	0,71	0,73	0,75	0,77	0,78
33	0,85	0,86	0,88	0,89	0,90	0,91	0,92	0,92	0,93	0,94	0,94	0,95
34	0,41	0,43	0,46	0,49	0,52	0,55	0,57	0,60	0,63	0,66	0,69	0,71
35	0,61	0,63	0,65	0,67	0,69	0,70	0,72	0,74	0,76	0,77	0,79	0,80
36	0,09	0,09	0,10	0,11	0,12	0,14	0,15	0,17	0,19	0,22	0,25	0,28

Fonte: Elaborado pela autora.

Tabela 4.13: Probabilidade de resposta correta para cada um dos itens considerados nos níveis selecionados pertencentes ao grupo 5.

Item	Níveis na escala - Grupo 5																	
	340	345	350	355	360	365	370	375	380	385	390	395	400	405	410	415	420	425
2	0,86	0,87	0,89	0,90	0,91	0,93	0,94	0,94	0,95	0,96	0,96	0,97	0,97	0,98	0,98	0,98	0,98	0,99
3	0,93	0,93	0,94	0,95	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99
4	0,97	0,97	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
5	0,41	0,42	0,44	0,46	0,48	0,50	0,53	0,55	0,57	0,60	0,62	0,64	0,67	0,69	0,71	0,73	0,75	0,77
6	0,87	0,88	0,89	0,90	0,91	0,92	0,93	0,94	0,94	0,95	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98
7	0,77	0,79	0,80	0,82	0,83	0,84	0,85	0,86	0,88	0,89	0,89	0,90	0,91	0,92	0,92	0,93	0,94	0,94
8	0,41	0,44	0,46	0,49	0,52	0,54	0,57	0,60	0,63	0,66	0,69	0,72	0,74	0,77	0,79	0,81	0,83	0,85
9	0,90	0,91	0,92	0,93	0,94	0,95	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99
10	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
11	0,94	0,94	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99
12	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
13	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00
14	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
15	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
16	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
17	0,53	0,56	0,60	0,63	0,66	0,69	0,72	0,75	0,77	0,80	0,82	0,84	0,86	0,88	0,89	0,90	0,92	0,93
18	0,97	0,98	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00
19	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
20	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
21	0,97	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
22	0,87	0,88	0,88	0,89	0,89	0,90	0,91	0,91	0,92	0,92	0,93	0,93	0,93	0,94	0,94	0,95	0,95	0,95
23	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00
24	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99	1,00	1,00	1,00
25	0,69	0,69	0,70	0,71	0,71	0,72	0,73	0,74	0,74	0,75	0,76	0,76	0,77	0,77	0,78	0,79	0,79	0,80
26	0,91	0,92	0,93	0,93	0,94	0,94	0,95	0,95	0,95	0,96	0,96	0,96	0,97	0,97	0,97	0,97	0,98	0,98
27	0,51	0,53	0,55	0,57	0,59	0,61	0,63	0,65	0,67	0,69	0,71	0,73	0,74	0,76	0,78	0,79	0,81	0,82
28	0,92	0,94	0,94	0,95	0,96	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	1,00
29	0,36	0,38	0,40	0,42	0,44	0,47	0,49	0,51	0,53	0,56	0,60	0,63	0,65	0,65	0,67	0,69	0,71	0,74
31	0,91	0,92	0,93	0,93	0,94	0,95	0,95	0,96	0,96	0,97	0,97	0,98	0,98	0,98	0,98	0,99	0,99	0,99
32	0,80	0,82	0,83	0,85	0,86	0,88	0,89	0,90	0,91	0,92	0,93	0,93	0,94	0,95	0,95	0,96	0,96	0,97
33	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,98	0,98	0,98	0,99	0,99	0,99	0,99	0,99	0,99	0,99
34	0,74	0,76	0,78	0,80	0,82	0,84	0,86	0,87	0,89	0,90	0,91	0,92	0,93	0,94	0,95	0,95	0,96	0,96
35	0,82	0,83	0,84	0,85	0,86	0,87	0,88	0,89	0,90	0,91	0,92	0,92	0,93	0,94	0,94	0,95	0,95	0,96
36	0,31	0,35	0,38	0,42	0,47	0,51	0,55	0,60	0,64	0,68	0,72	0,75	0,78	0,81	0,84	0,86	0,88	0,90

Fonte: Elaborado pela autora.

4.5.2 Interpretação da Escala

O próximo passo após a segmentação da escala é a interpretação pedagógica, realizada de acordo com os itens posicionados como âncora nos grupos, representando a divisão da escala. Para cada grupo foi atribuído um nome, como mostra a Tabela 4.14, com a finalidade de criar uma ideia de hierarquia e ordenamento e assim facilitar a interpretação.

Tabela 4.14: Nome atribuído para cada grupo de acordo com a divisão da escala.

Grupos	Faixa de Habilidade	Nome Atribuído
1	[75 a 155]	Insuficiente
2]155 a 230]	Abaixo do Básico
3]230 a 275]	Básico
4]275 a 335]	Acima do Básico
5]335 a 425]	Proficiente

Fonte: Elaborado pela autora.

Na fase de posicionamento dos itens, os correspondentes às competências 21, 22, 25, 26, 27 e 35 não foram posicionados como âncora, em nenhum nível da escala, por apresentarem valores para os parâmetros a abaixo de 1. Na sequência, verificou-se em que grupo estes itens estariam posicionados, caso tivessem sido classificados como âncora. Feito isso, procedeu-se a análise das competências abordadas, nos respectivos grupos, com base nos itens âncora e nos demais. Assim sendo, a competência abordada pelo item 22 foi alocada na interpretação do grupo 2 e a abordada pelo item 35 alocada na interpretação do grupo 4. Os itens 21, 25, 26 e 27 não foram considerados na interpretação, por não serem condizentes com os demais no grupo.

O quadro 4.2 traz a interpretação da escala de proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo e similares para cada faixa de proficiência. É importante salientar aqui que a interpretação da escala é cumulativa, ou seja, os alunos que estão situados em um determinado grupo dominam não só as habilidades associadas a esse grupo, mas também as proficiências descritas nos grupos anteriores.

Quadro 4.2: Descrição da escala.

Descrição das faixas de proficiência da escala
<p>Grupo 1 - Insuficiente [75 a 155]</p> <p>Os alunos com proficiência estimada nesta faixa, certamente:</p> <ul style="list-style-type: none"> • Resolvem problemas com números naturais, envolvendo operações de adição e multiplicação. • Reconhecem representações de um número racional. Reconhecem a equivalência entre números racionais, apresentados na forma fracionária e efetuam a divisão do numerador pelo denominador obtendo a representação na forma decimal. Entendem que este decimal representa a forma percentual do número racional, quando multiplicado por 100.
<p>Grupo 2 - Abaixo do Básico [155 a 230]</p> <p>Os alunos com proficiência estimada nesta faixa, certamente:</p> <ul style="list-style-type: none"> • Identificam fração como representação de uma situação problema. Compreendem a correspondência entre uma situação relatada e a fração que representa uma razão. Identificam o complementar de uma fração.

(continua)

(continuação)

- Identificam frações equivalentes. Reconhecem a representação de frações equivalentes através da simplificação dos numeradores e denominadores.
- Resolvem problemas com números racionais, envolvendo as operações de adição, multiplicação e divisão. Reconhecem a terça parte de um objeto, utilizando a operação de divisão e identificam a quantidade equivalente a uma fração desse objeto, utilizando a multiplicação ou a soma das partes calculadas.
- Substituem as variáveis de uma expressão por números inteiros e calculam seu valor numérico. Efetuando, também, o cálculo de potência e multiplicação entre números negativos.
- Reconhecem regularidade em uma sequência e a representam por meio de uma expressão algébrica. Identificam a correspondência entre duas variáveis e reconhecem a expressão algébrica que a representa.
- Reconhecem a expressão algébrica que representa uma função a partir de uma tabela. Identificam a relação expressa pelos dados fornecidos, observando a partir da tabela uma regularidade a cada linha.
- Identificam o gráfico que representa uma situação descrita em um texto.
- Resolvem problemas envolvendo modelagem, por meio de uma equação do 1º grau. Transpõem os dados numéricos apresentados para uma situação geral, adaptada uma equação do tipo $y = ax + b$, ou ainda, raciocinam aritmeticamente sobre os dados numéricos, realizando operações inversas para obter a resposta exigida.
- Identificam a representação algébrica de uma função exponencial crescente. Reconhecem uma função exponencial simples (genericamente dada por $y = a^x$ sendo $a > 0$) e identificam que para valores de a maiores que 1 a função é crescente.

Grupo 3 -Básico [230 a 275].

Os alunos com proficiência estimada nesta faixa, certamente:

- Identificam a equação de uma reta a partir de dois pontos dados. Calculam o coeficiente angular da reta e determinam a equação a partir da substituição de um dos pontos dados e do coeficiente encontrado utilizando a equação da reta tangente ou ainda, determinam a equação através da resolução de um sistema de equações.
- Identificam que o ponto de interseção de duas retas pode ser determinado algebricamente a partir da montagem de um sistema linear de duas equações.
- Resolvem problemas envolvendo cálculo de perímetro de figuras planas a partir de uma situação problema. Identificam e calculam o perímetro de uma figura.
- Identificam a localização de números reais na reta numérica. Fornecidos dois pontos em uma reta numérica, reconhecem que as abscissas desses dois pontos dados estão localizadas entre 0 e 1 e identificam que o produto de dois números menores que um é igual a um número menor que o menor deles.
- Identificam a localização de números racionais na reta numérica. Reconhecem a ordem de crescimento dos números racionais na reta e identificam corretamente os valores solicitados.
- Efetuam cálculos que envolvem operações de adição, subtração, multiplicação e potenciação com números racionais (em uma expressão algébrica do tipo $[(1+1/n)(1-1/n)]^2$ onde, n é um número natural qualquer $\neq 0$). Operam adequadamente soma e subtração de números racionais, assim como o cálculo da potência e do produto.
- Reconhecem a representação algébrica de uma função de primeiro grau dado o seu gráfico. Reconhecem a equação da reta dada na forma implícita e a determinam através da análise do coeficiente angular e linear.

(continua)

(continuação)

- Resolvem problemas que envolvem o ponto de máximo de uma função quadrática. Identificam o ponto de máximo desta função, a partir do cálculo das coordenadas do vértice, apresentado o seu gráfico e a equação da parábola. Reconhecem as coordenadas do vértice como ponto de máximo.
- Reconhecem o gráfico de uma função polinomial de primeiro grau. Identificam que o gráfico de uma função polinomial de primeiro grau é dado por uma reta crescente ou decrescente.

Grupo 4 - Acima do Básico [275 a 335].

Os alunos com proficiência estimada nesta faixa, certamente:

- Interpretam geometricamente os coeficientes da equação de uma reta. Reconhecem as propriedades da reta (dada na sua forma genérica como $y = ax + b$) manipulando a equação dada, apresentada de forma implícita.
- Resolvem problemas envolvendo cálculo de área de figuras planas. Identificam figuras geométricas formadas em uma figura fornecida. Efetuam o cálculo da área de triângulos, círculos e do quadrado, realizando operações de subtração entre as áreas encontradas e chegam ao resultado solicitado.
- Relacionam as raízes de um polinômio com a sua decomposição em fatores do primeiro grau.
- Reconhecem as propriedades da função logarítmica (do tipo $\log_a x$ com $a > 1$) e identificam a função a partir de sua representação gráfica.
- Resolvem problemas que envolvam função exponencial. Identificam a função exponencial que expressa uma situação descrita no texto. Estabelecem corretamente a relação entre o valor da imagem de uma função com seu valor correspondente no domínio dessa função. Utilizam as propriedades de potenciação necessárias para a resolução.

Grupo 5 - Proficiente [335 a 425].

Os alunos com proficiência estimada nesta faixa, certamente:

- Identificam, dentre as equações do segundo grau com duas incógnitas, a que representa uma circunferência. Reconhecem as propriedades de uma circunferência (genericamente, essa figura geométrica possui centro (x_c, y_c) e raio R). Reconhecem que como qualquer de seus pontos é equidistante do centro, a distância deles ao centro é o raio R . A partir dessas propriedades, identificam a equação reduzida da circunferência utilizando-se o teorema de Pitágoras, dada por $(x - x_c)^2 + (y - y_c)^2 = R^2$. A partir desta, chegam a equação geral da circunferência, que é dada por $x^2 + y^2 - 2xx_c - 2yy_c + x_c^2 + y_c^2 - R^2 = 0$.
- Resolvem problemas envolvendo o volume de sólidos. Calculam o volume total de um cilindro e o volume de esferas efetuando subtração e divisão entre esses volumes em uma situação problema.
- Efetuam cálculos com valores aproximados de radicais. Realizam cálculos que exigem técnicas de substituição, transformação de um número racional em uma fração decimal, propriedades de radiciação e potência.
- Resolvem equações modulares e encontram qual é o conjunto solução de uma equação modular (do tipo $|ax - b| = cx - d$ onde a, b, c e d são números naturais quaisquer). Identificam para quais valores da incógnita x a equação possui solução.
- Identificam a representação gráfica das funções trigonométricas e reconhecem suas propriedades. Distinguem corretamente a representação gráfica de funções (do tipo $y = a \sin bx$ com a e b positivos).

Considerações Finais e Perspectivas Futuras

Este trabalho consistiu em um estudo sobre a construção de escalas de proficiência com base na teoria da resposta ao item. Para isto, foram consideradas as escalas construídas para o SAEB, para o SARESP e proposta por Beaton e Allen (1992). Mais especificamente, consistiu na construção de uma escala para medir proficiência em conteúdos matemáticos básicos, aprimorando a escala elaborada inicialmente por Rossi (2015), a partir das escalas estudadas e propostas da literatura. Além disso, adequou-se o instrumento para as próximas aplicações da prova, via *web*, com a finalidade de agilizar a construção dos escores dos alunos e obtenção dos resultados.

Rossi (2015) trabalhou na construção de uma escala para medir proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo diferencial e integral e similares, de ingressantes em cursos da área de exatas. Para isto, com o auxílio de pesquisadores da área de cálculo e de avaliação educacional, construiu a matriz de referência, um banco de questões e elaborou uma prova que foi aplicada para ingressantes na FCT/UNESP nos cursos da área de exatas e para estudantes do terceiro ano do ensino médio, de uma escola pública da cidade de Presidente Prudente, em outubro de 2014. O modelo logístico unidimensional de três parâmetros foi utilizado para calibração (estimação dos parâmetros) dos itens e cálculo das proficiências, utilizando para isso a abordagem bayesiana. Os resultados do trabalho de Rossi (2015) mostraram que a maioria dos itens que compuseram a prova foram importantes para medir a proficiência desejada. No entanto, mostraram também que em parte da escala, novos itens deveriam ser incluídos.

Assim, considerando a necessidade de alterações em itens da prova aplicada em 2014, uma nova proposta foi construída, com base na mesma matriz de referência elaborada por Rossi (2015), sendo mantidos 26 dos 32 itens originais, modificados 5 deles, substituído 1 e contou ainda com a inclusão de mais 4 itens, totalizando para a nova prova 36 itens, sendo um para cada uma das competências que compõem a referida matriz.

Em 2017 todo o processo foi repetido com a nova proposta de prova, aplicada a todos os alunos matriculados no primeiro semestre do primeiro ano dos cursos da área de exatas da FCT/UNESP. Os parâmetros dos itens, foram novamente calibrados, inclusive dos itens mantidos na prova, já aplicados anteriormente, em 2014. Optou-se por criar a escala e interpretá-la a partir da prova atual, pois nesta aplicação todos os respondentes são alunos matriculados no primeiro semestre do primeiro ano de cada curso participante.

Assim, acredita-se que os resultados obtidos são mais adequados para a construção e interpretação da escala proposta.

Os resultados confirmaram a eficiência dos itens para estimar a proficiência dos respondentes. Em busca de uma adequada interpretação para escala, inicialmente optou-se por realizar o posicionamento dos itens como âncora, nos níveis selecionados para interpretação, utilizando os procedimentos adotados pelo SAEB, SARESP e o sugerido por Beaton e Allen (1992) utilizado por Rossi (2015). Porém, os resultados obtidos mostraram, nos três procedimentos utilizados, que em alguns dos níveis não foram posicionados itens âncora, o que dificultaria uma boa interpretação da escala em toda sua amplitude.

Diante disso, optou-se por adotar a proposta sugerida por Barros (2016) e segmentar a escala em grupos, ou seja, em faixas de proficiência utilizando a análise de agrupamentos hierárquicos. Desta forma, cada uma das faixas de proficiência pôde ser caracterizada de acordo com os itens âncora, posicionados em cada uma delas, a partir de suas probabilidades de resposta correta e dos valores do parâmetro de discriminação. Assim, foi possível construir a interpretação da escala em toda a sua amplitude.

Embora os resultados apresentados sejam consistentes, a curva da informação do teste mostrou que novos itens precisam ser incluídos para se obter mais informação nas extremidades da escala, o que evidência a necessidade de um processo contínuo de aprimoramento do banco de questões, e conseqüentemente, do aprimoramento desta escala de proficiência.

Para dar continuidade ao processo de aperfeiçoamento do banco de questões e aprimoramento da escala de proficiência, foi colocado em prática a proposta de aplicação da prova para este ano de 2018 (início do primeiro semestre), via *web*, a partir de um projeto paralelo em desenvolvimento pela Professora Dra Mariana Curi USP/ICMC, com a utilização do Sistema Moodle (*Modular Object Oriented Distance Learning*), que consiste em um sistema de gerenciamento para criação de cursos onlines. Para realizar a prova os alunos se cadastram como visitantes e se inscrevem em um curso de extensão, criado especificamente para este fim. Outra proposta de aplicação da prova via *web*, encontra-se em desenvolvimento na FCT/UNESP através de um projeto do Professor Dr. Mário H. Tarumoto.

A prova atual traz duas modificações em relação a prova aplicada em 2017, a substituição dos itens 1 e 25, por apresentarem baixo valor para o parâmetro de discriminação. Para a inserção das questões na plataforma contou-se com orientações da Professora Dra Mariana Curi e com a colaboração de uma aluna de Doutorado da USP/ICMC.

Novamente, a prova será aplicada aos alunos matriculados no primeiro semestre do primeiro ano nos cursos da área de exatas da FCT/UNESP e contará com a participação dos alunos ingressantes dos cursos de Matemática e Estatística da USP/ICMC. Após a aplicação da prova será realizada a equalização dos resultados em relação aos obtidos em 2017. Desta forma, com a equalização, as provas aplicadas estarão na mesma escala e os resultados obtidos nos anos de 2017 e 2018 poderão ser comparados.

A proposta para a construção da escala de proficiência em conteúdos matemáticos básicos, necessários para o acompanhamento das disciplinas de cálculo e similares, com base na TRI e sua interpretação são contribuições consideradas relevantes neste trabalho. Além disso, a participação na elaboração de propostas para adequação da prova, para aplicação via *web*, também configura contribuições.

Como perspectivas futuras de trabalho considera-se o teste de novos itens do banco de questões e, conseqüentemente, o aprimoramento constante da escala de proficiência com a aplicação da prova nos anos seguintes e a investigação da influência do meio utilizado para aplicação da prova (impressa ou via *web*) na proficiência do respondente. O acompanhamento e a comparação da proficiência dos alunos ao longo dos anos devem ser utilizados como ferramentas de apoio no direcionamento da aprendizagem.

Referências

- ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. da C. *Teoria de Resposta ao Item: Conceitos e Aplicações*. São Paulo: 4º SINAPE, ABE, 2000. 154 p.
- ANJOS, A. dos; ANDRADE, D. F. de. *Teoria de Resposta ao Item com o uso do R*. João Pessoa: 20º SINAPE, 2012. 104 p.
- AZEVEDO, C. L. N. *Métodos de Estimação na Teoria de Resposta ao Item, 2003, 121 f.* Dissertação (Mestrado) — Universidade de São Paulo, 2003.
- BAKER, F. B. *The Basics of Item Response Theory*. 2. ed. University of Wisconsin: ERIC Clearinghouse on Assessment and Evaluation, 2001. 174 p.
- BAKER, F. B.; KIM, S.-H. *Item response theory: Parameter estimation techniques*. 2. ed. Nova York: CRC Press, 2004. 419 p.
- BARROS, G. T. d. F. *Procedimentos para a construção de indicadores por meio da teoria de resposta ao item: a criação de uma medida de nível socioeconômico familiar*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, 2016.
- BEATON, A. E.; ALLEN, N. L. Chapter 6: Interpreting scales through scale anchoring. *Journal of Educational Statistics*, Sage Publications Sage CA: Los Angeles, CA, v. 17, n. 2, p. 191–204, 1992.
- BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: _____. *Estatistical Theories of Mental Test Scores*. [S.l.: s.n.], 1968. cap. 17-20, p. 397–547.
- BOCK, R. D.; AITKIN, M. Marginal maximum likelihood estimation of item parameters: Application of an em algorithm. *Psychometrika*, Springer, v. 46, n. 4, p. 443–459, 1981.
- BOCK, R. D.; LIEBERMAN, M. Fitting a response model for dichotomously scored items. *Psychometrika*, Springer, v. 35, n. 2, p. 179–197, 1970.
- BOCK, R. D.; ZIMOWSKI, M. F. Multiple group irt. In: *Handbook of modern item response theory*. [S.l.]: Springer, 1997. p. 433–448.
- BRASIL. INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. SAEB. 2017a. Disponível em: <<http://portal.inep.gov.br/web/guest/educacao-basica/saeb>>. Acesso em: 13 jun. 2017.
- BRASIL. MEC - Ministério da Educação. 2017b. Disponível em: <<http://portal.mec.gov.br/educacenso-sp-1181106924>>. Acesso em: 19 jun. 2017.
- BRASIL. MEC - Ministério da Educação. 2017c. Disponível em: <<http://portal.mec.gov.br/component/tags/tag/36188>>. Acesso em: 20 jun. 2017.

BRASIL. INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. ENEM. 2017d. Disponível em: <<http://portal.inep.gov.br/web/guest/enem>>. Acesso em: 26 jun. 2017.

BRASIL. Nota Técnica - INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. ENEM. 2017e. Disponível em: <http://download.inep.gov.br/educacao/_basica/enem/nota/_tecnica/2011/nota/_tecnica/_procedimento/_de/_calcul/_das/_notas/_enem/_2.pdf>. Acesso em: 17 jul. 2017.

BRASIL. Plano de Desenvolvimento da Educação PDE/SAEB 2011. 2017f. Disponível em: <http://portal.mec.gov.br/dmdocuments/saeb/_matriz2.pdf>. Acesso em: 20 abr. 2017.

BRASIL. Plano de Desenvolvimento da Educação PDE/Prova Brasil 2011. 2017g. Disponível em: <http://portal.mec.gov.br/dmdocuments/prova/%20brasil/_matriz2.pdf>. Acesso em: 20 abr. 2017.

CURTIS, S. M. et al. Bugs code for item response theory. *Journal of Statistical Software*, Foundation for Open Access Statistics, v. 36, n. 1, p. 1–34, 2010.

DANTE, L. R. *Matemática contexto e aplicações*. 1. ed. São Paulo: Ática, 2000. 590 p.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, JSTOR, p. 1–38, 1977.

DUTOIT, M. *IRT from SSI: BILOG-MG MULTILOG PARSCALE TESTFACT*. Scientific Software International, Inc., 2003. ISBN 089498053X,9780894980534. Disponível em: <<http://gen.lib.rus.ec/book/index.php?md5=A14F24300B79A508195AB8DCE0748251>>.

ESTEVES, A. S. Alves de L. *Aplicação da Teoria de Resposta ao Item para Avaliação da Proficiência em Conteúdos Matemáticos Básicos de Alunos Oriundos de Escolas Públicas*. Monografia (Trabalho de Conclusão de Curso) — Universidade Estadual Paulista, 2015.

FOX, J.-P. *Bayesian Item Response Modeling: Theory and Applications*. 1. ed. Springer-Verlag New York, 2010. (Statistics for Social and Behavioral Sciences). ISBN 1441907416,9781441907417. Disponível em: <<http://gen.lib.rus.ec/book/index.php?md5=54142BC4F8B3ED806CCFC31F40F032A7>>.

GAMERMAN, D.; LOPES, H. F. *Markov Chain Monte Carlos: Stochastic Simulation for Bayesian Inference*. 2. ed. Boca Raton: CRC Press, 2006. 323 p.

HAIR, J. F. et al. *Análise multivariada de dados*. [S.l.]: Bookman Editora, 2009.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. *Fundamentals of item response theory*. Califonia: Sage, 1991. 174 p.

KLEIN, R. Utilização da teoria de resposta ao item no sistema nacional de avaliação da educação básica (saeb). *Revista Meta: Avaliação*, v. 1, n. 2, p. 125–140, 2009.

LUNN, D. et al. The bugs project: Evolution, critique and future directions. *Statistics in medicine*, Wiley Online Library, v. 28, n. 25, p. 3049–3067, 2009.

- MARQUES, K. A. *Análise Bayesiana em Modelos TRI de três parâmetros*. 2008. 91 f. Dissertação (Mestrado) — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2008.
- MIGON, H. S.; GAMERMAN, D.; LOUZADA, F. *Statistical inference: an integrated approach*. 2. ed. Boca Raton: CRC press, 2014. 262 p.
- MISLEVY, R. J. Bayes modal estimation in item response models. *Psychometrika*, Springer, v. 51, n. 2, p. 177–195, 1986.
- MOREIRA JUNIOR, F. d. J. Aplicações da teoria da resposta ao item (tri) no brasil. *Rev. Bras. Biom*, v. 28, n. 4, p. 137–170, 2010.
- NOJOSA, R. T. Teoria da resposta ao item (tri): modelos multidimensionais. *Estudos em Avaliação Educacional*, n. 25, p. 123–166, 2002.
- PASQUALI, L. *Psicometria. Teoria dos testes na Psicologia e na Educação*. 2. ed. Petrópolis: Vozes, 2003. 397 p.
- ROSSI, P. R. *Construção de uma Escala para Avaliação da Proficiência em Conteúdos Matemáticos Básicos*. Dissertação (Mestrado) — Universidade Estadual Paulista, 2015.
- SARESP. *Relatório Técnico Estatístico Saresp - 2017 versão 01*. [S.l.: s.n.], 2017.
- SÃO PAULO. *Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo - (SARESP)*. 2017b. Disponível em: <<http://saresp.vunesp.com.br/resultados3em.html>>. Acesso em: 30 out. 2017.
- TEAM, R. *RStudio: Integrated Development Environment for R*. Boston, MA, 2015. Disponível em: <<http://www.rstudio.com/>>.
- VALLE, R. da C. Teoria de resposta ao item. *Estudos em avaliação educacional*, n. 21, p. 7–92, 2000.

Instrumento de Avaliação

Projeto: Construção de Escalas para Avaliação de Proficiência em Conteúdos Matemáticos Básicos com Base na Teoria da Resposta ao Item.



Nome do aluno: _____
RG: _____

LEIA ATENTAMENTE AS INSTRUÇÕES SEGUINTE

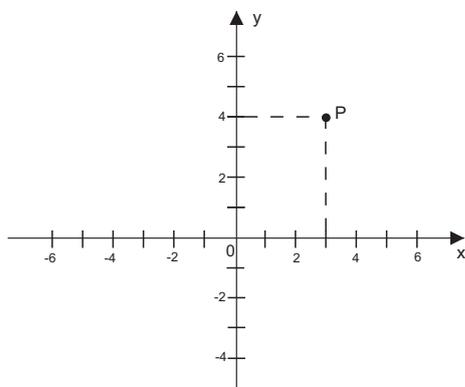
1. Preencha o cartão resposta com seu nome completo, assinatura e documento de identidade (observe que idade e sexo já fazem parte das respostas).
 2. Cada questão tem cinco alternativas de resposta: (A), (B), (C), (D) e (E) e apenas uma delas é correta.
 3. Para cada questão marque a alternativa escolhida no cartão-resposta, preenchendo todo o espaço dentro do retângulo correspondente com caneta esferográfica azul ou preta.
 4. Marque apenas uma alternativa para cada questão.
- Atenção:** questão com mais de uma alternativa marcada será anulada, mesmo que uma das marcações esteja correta.
5. Não é permitido o uso de instrumentos de desenho, calculadoras, aparelhos eletrônicos ou quaisquer fontes de consulta.
 6. Os espaços em branco na prova podem ser usados para rascunho.
 7. Ao final, a prova deverá ser entregue juntamente com o cartão-resposta.

É com grande alegria que contamos com a sua participação neste projeto. Encare as questões desta prova como quebra-cabeças interessantes e divirta-se com a busca de suas soluções.

desejamos que faça uma boa prova!

Questão 1

A figura abaixo mostra um ponto em um plano cartesiano.



As coordenadas do ponto P são:

- (A) (6, 6).
- (B) (-3, 4).
- (C) (3, 4).
- (D) (3, 7).
- (E) (4, 5).

Questão 2

A reta da equação $2y + x = 0$:

- (A) é paralela ao eixo OX.
- (B) é paralela ao eixo OY.
- (C) tem coeficiente angular $-\frac{1}{2}$.
- (D) tem coeficiente angular $\frac{1}{2}$.
- (E) tem coeficiente angular 2.

Questão 3

Qual é a equação da reta que contém os pontos (3, 5) e (4, -2)?

- (A) $y = -7x + 26$.
- (B) $y = -\frac{1}{7}x - \frac{10}{7}$.
- (C) $y = \frac{1}{7}x - \frac{18}{7}$.
- (D) $y = x + 2$.
- (E) $y = 7x - 16$.

Questão 4

O ponto de interseção das retas de equações $x + 3y - 1 = 0$ e $x - y + 3 = 0$ é:

- (A) (1, -2).
- (B) (-2, 1).
- (C) (-1, -2).
- (D) (-2, -1).
- (E) (4, 1).

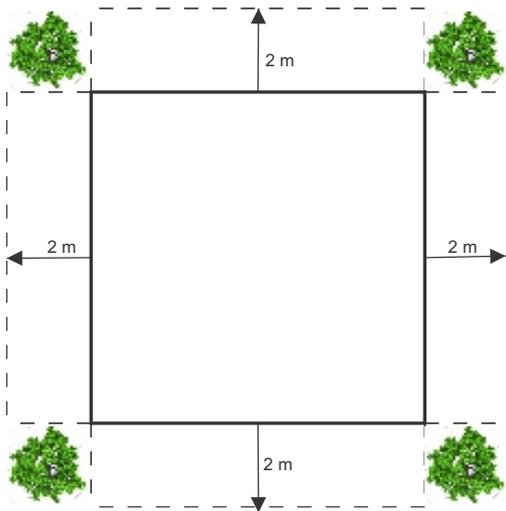
Questão 5

Dentre as equações abaixo, pode-se afirmar que a de uma circunferência é:

- (A) $(x - 1)^2 + y^2 = 25$.
- (B) $x^2 - y - 4x = -3$.
- (C) $x^2 + y^2 = -16$.
- (D) $x^2 - y - 9 = 0$.
- (E) $x^2 - y^2 - 4x = 9$.

Questão 6

Uma praça quadrada, que possui o perímetro de 24 metros, tem uma árvore próxima de cada vértice e fora dela. Deseja-se aumentar a área da praça, alterando-se sua forma e mantendo as árvores externas a ela, conforme ilustra a figura.

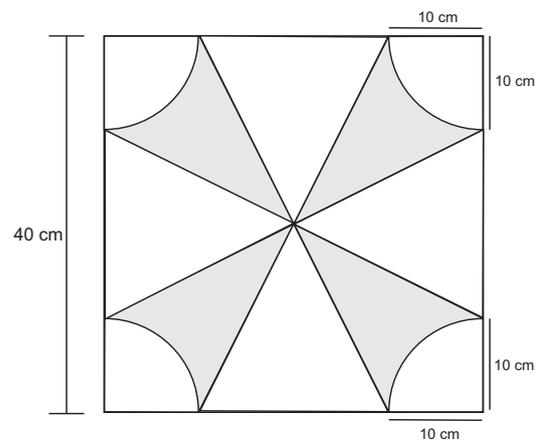


O novo perímetro da praça é:

- (A) 24 metros.
- (B) 32 metros.
- (C) 36 metros.
- (D) 40 metros.
- (E) 64 metros.

Questão 7

Paulo resolve modificar o revestimento do piso de sua sala de estar e escolhe uma cerâmica cujo formato está representado na figura a seguir. A cerâmica escolhida tem a forma de um quadrado cujo o lado mede 40 cm e possui 4 arcos de circunferência, de raio igual a 10 cm, cujos centros estão localizados nos vértices do quadrado.



Com base nessas informações, qual é a área do desenho formado na cerâmica, em centímetros quadrados? (Considere $\pi = 3,14$)

- (A) 314.
- (B) 400.
- (C) 486.
- (D) 1114.
- (E) 1286.

Questão 8

Um corpo cilíndrico, com 4 cm de raio e 12 cm de altura, está com água até a altura de 8 cm. Foram colocadas em seu interior bolas de gude de 2 cm de diâmetro, e o nível da água atingiu a boca do vidro, sem derramamento.

Quantas bolas de gude foram colocadas?

- (A) 32.
- (B) 48.
- (C) 64.
- (D) 80.
- (E) 96.

Questão 9

Na figura abaixo estão representados os números reais 0, x, y, 1.



A posição do produto $x \times y$ é:

- (A) à esquerda do zero.
- (B) entre 0 e x.
- (C) entre x e y.
- (D) entre y e 1.
- (E) à direita de 1.

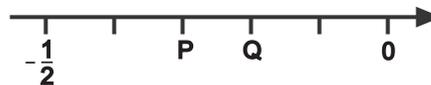
Questão 10

Num cinema, há 12 fileiras com 16 poltronas e 15 fileiras com 18 poltronas. O número de poltronas é:

- (A) 78.
- (B) 192.
- (C) 270.
- (D) 462.
- (E) 480.

Questão 11

A figura abaixo mostra os pontos P e Q que correspondem a números racionais e foram posicionados na reta numerada do conjunto dos racionais.



Os valores atribuídos a P e Q, conforme suas posições na reta numérica são:

- (A) $P = -\frac{1}{5}$ e $Q = -\frac{3}{10}$.
- (B) $P = -\frac{3}{10}$ e $Q = -\frac{1}{5}$.
- (C) $P = -\frac{3}{5}$ e $Q = -\frac{7}{10}$.
- (D) $P = -\frac{7}{10}$ e $Q = -\frac{3}{5}$.
- (E) $P = -\frac{3}{5}$ e $Q = -\frac{1}{5}$.

Questão 12

No Brasil $\frac{3}{4}$ da população vive na zona urbana.

De que outra forma podemos representar esta fração?

- (A) 7,5%.
- (B) 15%.
- (C) 25%.
- (D) 34%.
- (E) 75%.

Questão 13

Dos 11 jogadores de um time de futebol, apenas 5 tem menos de 25 anos de idade.

A fração de jogadores com mais de 25 anos de idade é:

- (A) $\frac{5}{6}$.
- (B) $\frac{6}{5}$.
- (C) $\frac{5}{11}$.
- (D) $\frac{6}{11}$.
- (E) $\frac{5}{25}$.

Questão 14

Quatro amigos, João, Pedro, Ana e Maria saíram juntos para fazer um passeio por um mesmo caminho. Até agora, João andou $\frac{6}{8}$ do caminho; Pedro, $\frac{9}{12}$; Ana, $\frac{3}{8}$ e Maria, $\frac{4}{6}$. Os amigos que se encontraram no mesmo ponto do caminho são:

- (A) João e Pedro.
- (B) João e Ana.
- (C) Ana e Maria.
- (D) Pedro e Ana.
- (E) Maria e João.

Questão 15

A professora de matemática propôs como exercício a expressão:

$$\left[\left(1 + \frac{1}{3} \right) \left(1 - \frac{1}{3} \right) \right]^2$$

Os alunos que resolveram corretamente a expressão encontraram como resultado:

- (A) $-\frac{8}{9}$.
- (B) 0.
- (C) $\frac{8}{9}$.
- (D) 2.
- (E) $\frac{64}{81}$.

Questão 16

Uma horta comunitária será criada em uma área de 5100 m^2 . Para o cultivo de hortaliças, serão destinados $\frac{2}{3}$ dessa área.

Quantos metros quadrados serão utilizados nesse cultivo?

- (A) 170.
- (B) 340.
- (C) 1700.
- (D) 2550.
- (E) 3400.

Questão 17

Supondo $\sqrt[4]{8} = 1,68$, o valor mais próximo de $\sqrt{\frac{0,09}{\sqrt{2}}}$ é:

- (A) 25, 2.
- (B) 0, 0252.
- (C) 0, 252.
- (D) 2, 5.
- (E) 0, 00252.

Questão 18

O resultado da expressão $2x^2 - 3x + 10$, para $x = -2$, é:

- (A) -4.
- (B) 0.
- (C) 12.
- (D) 13.
- (E) 24.

Questão 19

As variáveis n e P assumem valores conforme mostra o quadro abaixo.

n	5	6	7	8	9	10
P	8	10	12	14	16	18

A relação entre P e n é dada pela expressão:

- (A) $P = n + 1$.
- (B) $P = n + 2$.
- (C) $P = 2n - 2$.
- (D) $P = n - 2$.
- (E) $P = 2n$.

Questão 20

Uma empresa, em processo de reestruturação, propôs a seus funcionários, admitidos há pelo menos dois anos, uma indenização financeira para os que pedissem demissão, que varia em função do número de anos trabalhados. A tabela abaixo era utilizada para calcular o valor (i) da indenização, em função do tempo trabalhado (t).

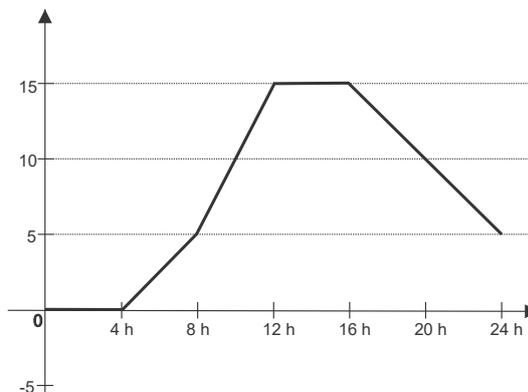
Tempo trabalhado (em anos)	Valor da indenização (em reais)
1	450
2	950
3	1450
4	1950

A expressão que permite determinar o valor de indenização i para t anos trabalhados é:

- (A) $i = 450t$.
 (B) $i = 450 + 500t$.
 (C) $i = 450(t - 1)$.
 (D) $i = 450 + 500(t - 1)$.
 (E) $i = 500t$.

Questão 21

O gráfico abaixo mostra a temperatura numa cidade da Região Sul, em um dia do mês de julho.



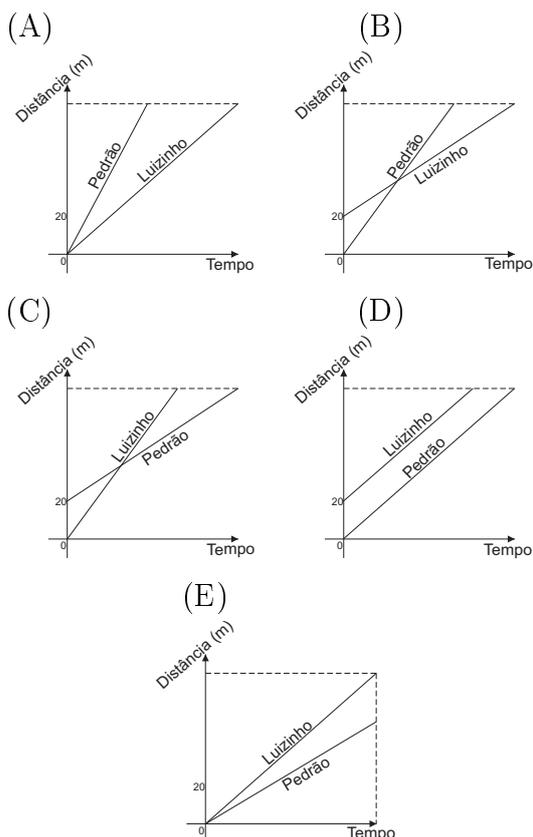
De acordo com o gráfico, a temperatura aumenta no período de:

- (A) 8 às 16h.
 (B) 16 às 24h.
 (C) 4 às 12h.
 (D) 12 às 16h.
 (E) 4 às 16h.

Questão 22

Luizinho desafia seu irmão mais velho, Pedrão, para uma corrida. Pedrão aceita e permite que o desafiante saia 20 metros a sua frente. Pedrão ultrapassa Luizinho e ganha a corrida.

O gráfico que melhor ilustra essa disputa é:



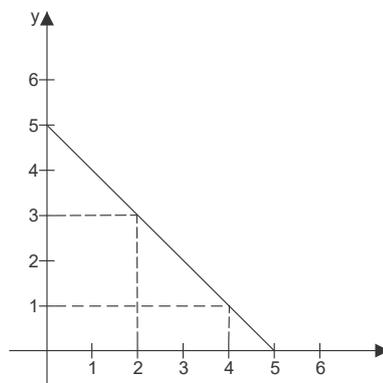
Questão 23

O custo de produção de uma pequena empresa é composto por um valor fixo de R\$ 1.500,00 mais R\$ 10,00 por peça fabricada. O número x de peças fabricadas quando o custo é de R\$ 3.200,00 é:

- (A) 150.
- (B) 160.
- (C) 170.
- (D) 320.
- (E) 470.

Questão 24

O gráfico abaixo mostra uma reta em um plano cartesiano.



Qual é a equação da reta apresentada no gráfico?

- (A) $x - y - 5 = 0$.
- (B) $x + y - 5 = 0$.
- (C) $x + y + 5 = 0$.
- (D) $x + y - 4 = 0$.
- (E) $x + y = 6$.

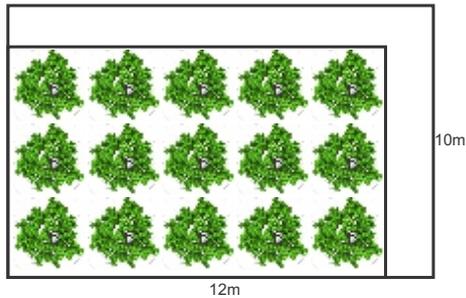
Questão 25

Por uma mensagem dos Estados Unidos para o Brasil, via fax, a Empresa de Correios e Telégrafos (ECT) cobra R\$ 1,37 pela primeira página e R\$ 0,67 por página que se segue, completa ou não. Qual o número mínimo de páginas de uma dessas mensagens para que seu preço ultrapasse o valor de R\$ 10,00?

- (A) 8.
- (B) 10.
- (C) 12.
- (D) 14.
- (E) 16.

Questão 26

Em um terreno retangular de tamanho $10\text{m} \times 12\text{m}$, deseja-se construir um jardim com 80 m^2 de área, deixando uma faixa para o caminho (sempre de mesma largura), como mostra a figura.



A largura do caminho deve ser de:

- (A) 1 metro.
- (B) 1,5 metros.
- (C) 2 metros.
- (D) 2,5 metros.
- (E) 3 metros.

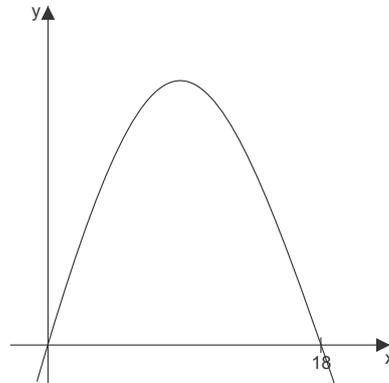
Questão 27

Seja M o conjunto dos números naturais n , tal que $2n^2 - 75n + 700 \leq 0$. Assim, é correto afirmar que:

- (A) apenas um dos elementos de M é múltiplo de 4.
- (B) apenas dois dos elementos de M são primos.
- (C) a soma de todos os elementos de M é igual a 79.
- (D) M contém exatamente 6 elementos.
- (E) os elementos de M são múltiplos de 3.

Questão 28

Uma bala é atirada de um canhão e sua trajetória descreve uma parábola de equação $y = -5x^2 + 90x$, onde as variáveis x e y são medidas em metros.



Nessas condições, a altura máxima atingida pela bala é:

- (A) 30,0 metros.
- (B) 40,5 metros.
- (C) 81,5 metros.
- (D) 405 metros.
- (E) 810 metros.

Questão 29

O conjunto solução da equação $|3x - 5| = 5x - 1$ é:

- (A) $\{-2\}$.
- (B) $\{\frac{3}{4}\}$.
- (C) $\{\frac{1}{5}\}$.
- (D) $\{2\}$.
- (E) $\{\frac{3}{4}, -2\}$.

Questão 30

São dados os conjuntos

$$A = \{x \in \mathbb{Z} \mid |x - 1| < 3\} \text{ e}$$

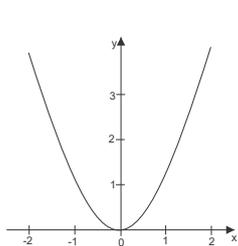
$$B = \{x \in \mathbb{Q} \mid |x| < 4\}$$

É correto afirmar que:

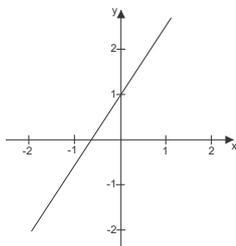
- (A) $A \cap B = \emptyset$.
- (B) $A \cup B = B$.
- (C) $A \subset B$ e $A \neq B$.
- (D) $A \cap B$ tem dois elementos.
- (E) $A \cup B$ tem infinitos elementos.

Questão 31

Observe os gráficos



(1)



(2)

Qual dos gráficos representa uma função polinomial de primeiro grau?

- (A) o gráfico (2).
- (B) o gráfico (1).
- (C) os dois gráficos.
- (D) o gráfico (2) apenas considerando a parte positiva.
- (E) Nenhum dos gráficos.

Questão 32

As raízes do polinômio $P(x) = (x - 3)^2(x + 1)$ são:

- (A) -2 e 1 .
- (B) 3 e -1 .
- (C) -3 e 1 .
- (D) 3 e 1 .
- (E) -3 e -1 .

Questão 33

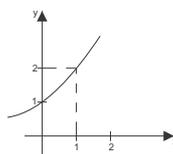
Abaixo estão relacionadas algumas funções.

Entre elas, a função exponencial crescente é:

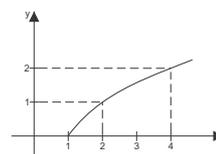
- (A) $f(x) = 5^{-x}$
- (B) $f(x) = \left(\frac{\sqrt{3}}{2}\right)^2$
- (C) $f(x) = (0, 1)^x$
- (D) $f(x) = 10^x$
- (E) $f(x) = 0, 5^x$

Questão 34

Nos gráficos (1) e (2) abaixo estão representadas duas funções.



(1)



(2)

Pode-se afirmar que:

- (A) $y = 2x$ está representada no (1).
- (B) $y = x^2 + 1$ está representada no (2).
- (C) $y = \log_2 x$ está representada no (2).
- (D) $y = 2^x$ está representada no (2).
- (E) $y = \log x$ está representada no (2).

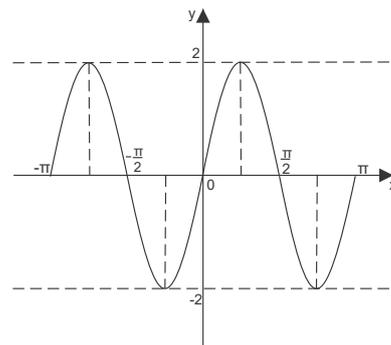
Questão 35

A população de bactérias em um meio pode ser modelada, sob certas condições, por uma função exponencial. Sabendo que a cada hora esta população duplica e que no tempo $t = 0$ existem 100 bactérias, a função $P(t)$ que representa a população em função do tempo t (em horas) é:

- (A) $p(t) = 100 \times 2^t$.
- (B) $p(t) = 100 + 2t$.
- (C) $p(t) = 2^t + 100$.
- (D) $p(t) = 2t$.
- (E) $p(t) = 2^t$.

Questão 36

O gráfico abaixo representa a função:



- (A) $y = -2 \cos x$.
- (B) $y = \cos \frac{x}{2}$.
- (C) $y = 2 \sin x$.
- (D) $y = \sin \frac{x}{2}$.
- (E) $y = 2 \sin 2x$.

Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Título da Pesquisa: “Avaliação dos Ingressantes nos Cursos da Área de Exatas na Proficiência em Conteúdos Matemáticos Básicos, Necessários para o Acompanhamento das Disciplinas de Cálculo e Similares”

Pesquisadora: Tânia Robaskiewicz Coneglian Fujii

Orientadora: Aparecida Donizete Pires de Souza

1. Natureza da pesquisa: Você está sendo convidado (a) a participar desta pesquisa que tem como finalidade a construção e interpretação de uma escala para medir o conhecimento em conteúdos matemáticos básicos, necessários para o acompanhamento em disciplinas de cálculo e similares com base na Teoria da Resposta ao Item, no nível de ensino médio.

2. Participantes da pesquisa: Serão convidados a participar desta pesquisa neste período, todos os alunos ingressantes nos cursos da área de exatas da FCT/Unesp do ano letivo de 2017.

3. Envolvimento na pesquisa: Participar como voluntário deste estudo contribuirá com as investigações propostas pela pesquisadora.

4. Riscos e desconforto: a participação nesta pesquisa não infringe as normas legais e éticas. Os procedimentos adotados nesta pesquisa obedecem aos Critérios da Ética em Pesquisa com Seres Humanos conforme Resolução no. 466/2012 do Conselho Nacional de Saúde. Nenhum dos procedimentos usados oferece riscos à sua dignidade.

5. Confidencialidade: todas as informações coletadas neste estudo são estritamente confidenciais. Somente a pesquisadora e sua orientadora (e/ou equipe de pesquisa) terão conhecimento de sua identidade e nos comprometemos a mantê-la em sigilo ao publicar os resultados dessa pesquisa.

6. Benefícios: ao participar desta pesquisa você não terá nenhum benefício direto. Entretanto, esperamos que este estudo traga informações importantes para produção de conhecimento científico, de forma a contribuir para a proposição de atividades que auxiliem os alunos no acompanhamento das disciplinas de cálculo e similares, onde a pesquisadora se compromete a divulgar os resultados obtidos, respeitando-se o sigilo das informações coletadas, conforme previsto no item anterior.

7. Pagamento: você não terá nenhum tipo de despesa para participar desta pesquisa, bem como nada será pago por sua participação. Você tem liberdade de se recusar a participar e ainda se recusar a continuar participando em qualquer fase da pesquisa, sem que tenha qualquer prejuízo. Sempre que quiser poderá pedir mais informações sobre a pesquisa através do telefone da pesquisadora do projeto e, se necessário através do telefone do Comitê de Ética em Pesquisa.

Após estes esclarecimentos, solicitamos o seu consentimento de forma livre para participar desta pesquisa. Portanto preencha, por favor, os itens que se seguem: Confiro que recebi uma via deste termo de consentimento, e autorizo a execução do trabalho de pesquisa e a divulgação dos dados obtidos neste estudo.

Obs: Não assine esse termo se ainda tiver dúvida a respeito.

Consentimento Livre e Esclarecido

Tendo em vista os itens acima apresentados, eu, de forma livre e esclarecida, manifesto meu consentimento em participar da pesquisa

Nome do Participante da Pesquisa

Assinatura do Participante da Pesquisa

Tânia Robaskiewicz Coneglian Fujii

Aparecida Donizete Pires de Souza

Pesquisadora: Tânia Robaskiewicz Coneglian Fujii Tel: (44) 99943-3081

Orientadora: Aparecida Donizete Pires de Souza Tel: (18) 3229-5617

Coordenadora do Comitê de Ética em Pesquisa: Profa. Dra. Edna Maria do Carmo

Vice-Coodenadora: Profa. Dra. Andreia Cristiane Silva Wiezzel

Telefone do Comitê: 3229-5315 ou 3229-5526

E-mail cep@fct.unesp.br

Proficiência dos Indivíduos em função do Número de Acertos

Tabela C.1: Número de acertos e proficiência por indivíduo na escala (250,50).

Indivíduo	Acertos	Habilidade	Indivíduo	Acertos	Habilidade
124	7	167,1	207	12	166,2
240	7	161,1	235	12	189,3
49	8	158,4	236	12	167,8
242	8	145,2	252	12	186,8
244	8	162,0	259	12	186,4
266	8	154,5	263	12	194,8
270	8	158,3	69	13	183,8
102	9	159,2	103	13	204,4
154	9	160,5	164	13	190,1
212	9	155,5	181	13	191,4
219	9	143,9	197	13	208,5
81	10	163,4	223	13	207,5
156	10	162,1	256	13	200,0
265	10	155,8	274	13	189,6
267	10	158,6	71	14	207,5
268	10	183,2	94	14	198,9
162	11	193,1	115	14	211,2
172	11	184,0	122	14	207,5
191	11	183,2	160	14	207,8
217	11	170,9	200	14	203,2
262	11	183,3	211	14	192,4
78	12	187,8	228	14	188,8
85	12	184,1	9	15	211,1
87	12	190,2	57	15	208,7
91	12	183,0	63	15	205,5
175	12	169,2	68	15	211,8
186	12	202,0	70	15	199,9

(continua)

			(continuação)		
Indivíduo	Acertos	Habilidade	Indivíduo	Acertos	Habilidade
92	15	222,5	56	18	232,6
134	15	219,1	66	18	221,7
166	15	209,8	77	18	223,4
205	15	194,6	88	18	235,5
209	15	224,5	98	18	242,5
218	15	204,9	113	18	240,8
227	15	184,3	155	18	232,7
241	15	206,9	168	18	253,7
254	15	221,7	188	18	242,4
31	16	233,4	192	18	236,5
34	16	218,8	250	18	238,1
64	16	222,1	258	18	236,5
76	16	212,9	271	18	240,6
90	16	223,7	272	18	237,3
93	16	222,1	19	19	234,3
108	16	223,8	40	19	234,1
126	16	216,2	75	19	248,7
129	16	216,8	99	19	238,7
167	16	217,0	111	19	239,8
169	16	207,2	132	19	232,6
176	16	209,9	138	19	243,1
210	16	213,9	152	19	248,3
222	16	194,5	177	19	241,3
238	16	226,9	180	19	242,4
2	17	231,6	185	19	235,3
4	17	223,7	189	19	248,5
17	17	234,1	196	19	232,0
35	17	235,2	216	19	217,5
41	17	229,8	260	19	244,6
65	17	222,8	276	19	241,5
67	17	211,8	37	20	246,5
72	17	221,2	44	20	259,2
109	17	205,9	50	20	245,1
136	17	216,9	59	20	251,8
137	17	235,2	73	20	260,0
141	17	214,1	83	20	249,9
157	17	234,7	89	20	245,2
182	17	230,3	110	20	252,8
198	17	213,3	114	20	236,9
214	17	218,1	118	20	251,0
224	17	237,4	121	20	246,9
226	17	214,0	128	20	252,7
264	17	239,9	187	20	257,8
33	18	240,9	206	20	246,2
46	18	232,7	230	20	248,8
47	18	248,8	232	20	256,9
48	18	248,3	234	20	251,8

(continua)

(continuação)

Indivíduo	Acertos	Habilidade	Indivíduo	Acertos	Habilidade
243	20	241,5	158	23	264,9
12	21	263,2	190	23	289,0
21	21	260,8	201	23	280,0
58	21	261,0	204	23	267,1
79	21	262,7	246	23	282,9
96	21	243,3	261	23	275,2
116	21	259,6	1	24	281,4
127	21	256,7	14	24	286,1
131	21	244,9	36	24	283,4
159	21	273,8	39	24	269,2
213	21	238,5	52	24	271,2
253	21	270,9	55	24	268,9
273	21	253,1	60	24	257,8
62	22	271,8	74	24	284,7
112	22	254,2	84	24	289,5
130	22	262,5	95	24	284,9
135	22	270,2	120	24	279,2
149	22	258,7	125	24	277,5
151	22	268,2	153	24	286,8
165	22	262,7	173	24	284,1
174	22	275,9	202	24	282,5
178	22	263,7	220	24	278,6
194	22	271,2	247	24	282,7
208	22	241,7	255	24	275,5
225	22	272,8	23	25	299,0
229	22	254,4	25	25	291,5
233	22	271,5	101	25	290,8
248	22	263,3	104	25	299,7
249	22	264,3	142	25	298,2
277	22	256,8	143	25	282,3
7	23	277,2	146	25	289,2
16	23	276,4	171	25	287,6
22	23	277,9	193	25	293,5
26	23	284,0	215	25	292,1
38	23	282,0	221	25	285,5
80	23	280,3	239	25	299,9
82	23	276,2	269	25	282,6
86	23	261,1	275	25	294,3
97	23	280,3	3	26	297,5
100	23	271,3	5	26	304,8
105	23	271,2	11	26	309,5
106	23	288,8	20	26	292,6
117	23	270,7	24	26	302,5
144	23	277,9	27	26	300,6
145	23	281,6	29	26	309,1
148	23	278,4	54	26	310,3
150	23	283,5	139	26	297,0

(continua)

(continuação)

Indivíduo	Acertos	Habilidade	Indivíduo	Acertos	Habilidade
179	26	298,9	251	28	325,8
6	27	312,1	15	29	332,3
51	27	308,1	28	29	326,8
119	27	301,9	53	29	323,5
195	27	308,9	123	29	331,2
18	28	319,3	140	29	332,5
43	28	316,6	161	29	330,5
45	28	320,4	257	29	325,1
61	28	307,0	8	30	341,1
107	28	316,7	10	30	332,5
133	28	304,0	170	30	340,3
147	28	328,7	199	30	337,9
163	28	318,7	13	32	368,4
183	28	320,2	30	32	356,6
184	28	314,0	42	32	371,1
203	28	326,2	237	32	367,9
231	28	317,1	32	34	392,1
245	28	314,7			

Curva Característica e Curva de Informação de cada Item

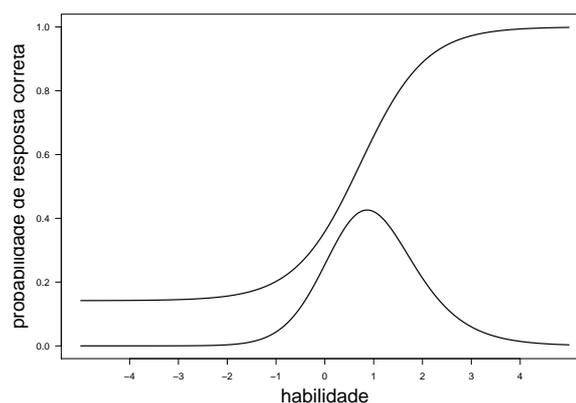
Item 2:

Tabela D.1: Gabarito, percentuais de resposta e parâmetros do item 2.

Indicadores	Percentuais de acertos					Parâmetros TRI		
Gabarito	A	B	C	D	E	a	b	c
C	10,11	8,66	41,88	19,49	19,13	1,49600	0,72680	0,14180

Fonte: Elaborado pela autora.

Figura D.1: Curva característica e curva de informação do item 2



Fonte: Elaborado pela autora.

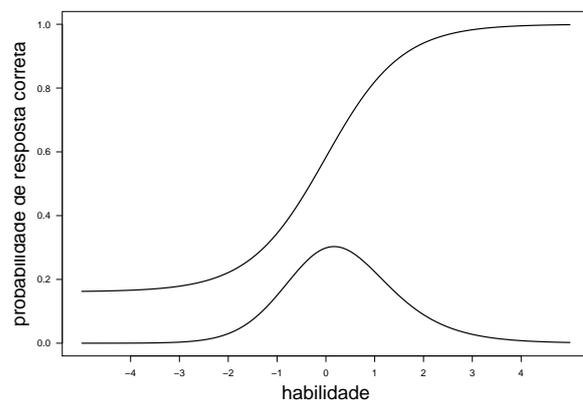
Item 3:

Tabela D.2: Gabarito, percentuais de resposta e parâmetros do item 3.

Indicadores	Percentuais de acertos					Parâmetros TRI		
Gabarito	A	B	C	D	E	a	b	c
A	58,12	11,55	6,86	6,50	16,97	1,28500	-0,00850	0,16120

Fonte: Elaborado pela autora.

Figura D.2: Curva característica e curva de informação do item 3



Fonte: Elaborado pela autora.

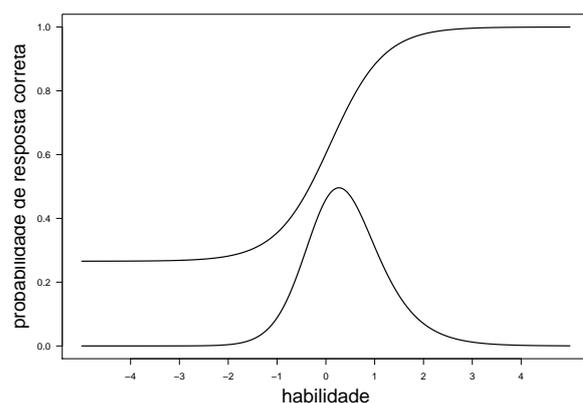
Item 4:

Tabela D.3: Gabarito, percentuais de resposta e parâmetros do item 4.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
B	8,66	61,73	10,47	10,47	8,30	1,81800	0,08940	0,26570

Fonte: Elaborado pela autora.

Figura D.3: Curva característica e curva de informação do item 4



Fonte: Elaborado pela autora.

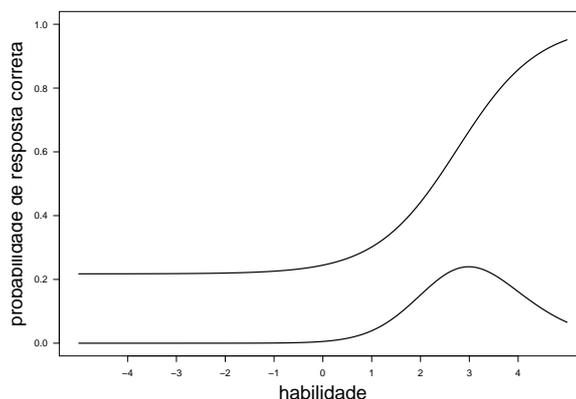
Item 5:

Tabela D.4: Gabarito, percentuais de resposta e parâmetros do item 5.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
A	29,96	8,30	28,88	16,25	15,88	1,20600	2,75400	0,21740

Fonte: Elaborado pela autora.

Figura D.4: Curva característica e curva de informação do item 5



Fonte: Elaborado pela autora.

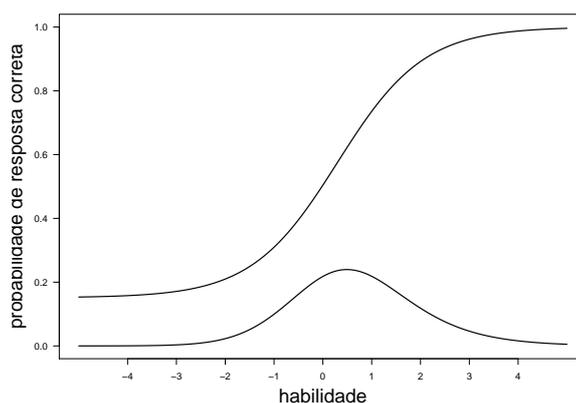
Item 6:

Tabela D.5: Gabarito, percentuais de resposta e parâmetros do item 6.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
D	3,61	35,38	4,33	51,99	4,69	1,13200	0,30040	0,15130

Fonte: Elaborado pela autora.

Figura D.5: Curva característica e curva de informação do item 6



Fonte: Elaborado pela autora.

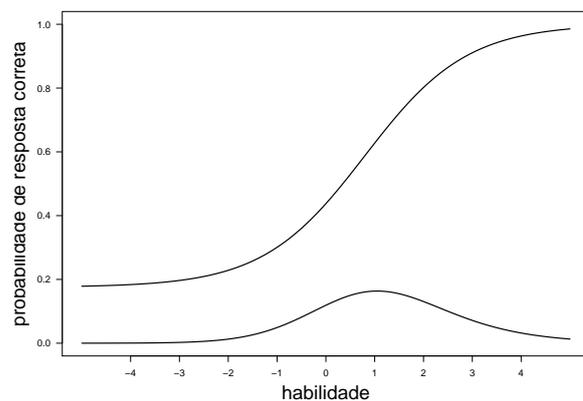
Item 7:

Tabela D.6: Gabarito, percentuais de resposta e parâmetros do item 7.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
C	10,83	9,75	46,93	18,41	13,36	0,95690	0,79820	0,17550

Fonte: Elaborado pela autora.

Figura D.6: Curva característica e curva de informação do item 7



Fonte: Elaborado pela autora.

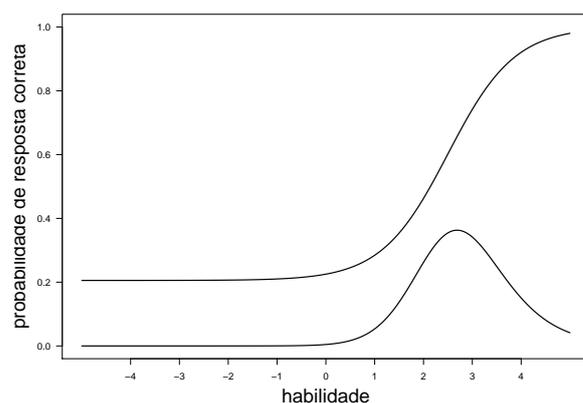
Item 8:

Tabela D.7: Gabarito, percentuais de resposta e parâmetros do item 8.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
B	38,99	27,44	17,69	6,86	8,30	1,46700	2,50200	0,20540

Fonte: Elaborado pela autora.

Figura D.7: Curva característica e curva de informação do item 8



Fonte: Elaborado pela autora.

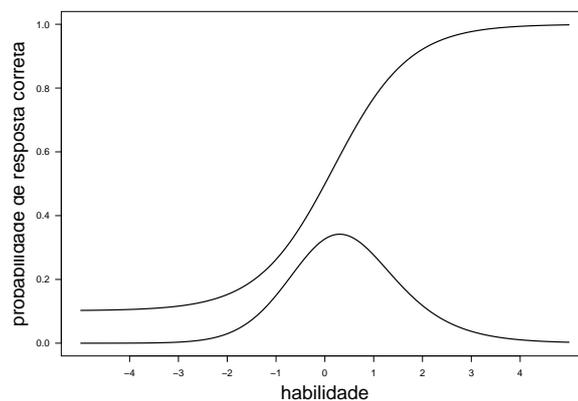
Item 9:

Tabela D.8: Gabarito, percentuais de resposta e parâmetros do item 9.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
B	2,53	50,90	24,55	11,19	10,83	1,29000	0,18030	0,10170

Fonte: Elaborado pela autora.

Figura D.8: Curva Característica e curva de informação do item 9



Fonte: Elaborado pela autora.

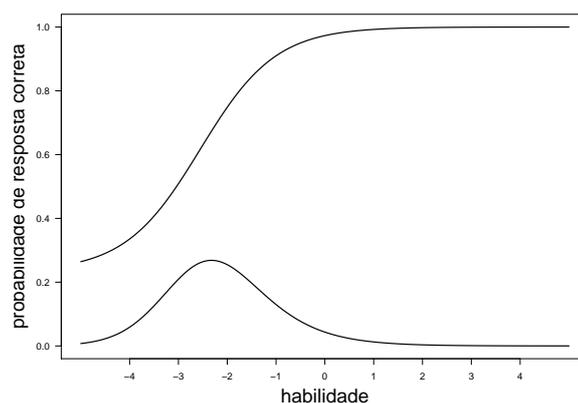
Item 10:

Tabela D.9: Gabarito, percentuais de resposta e parâmetros do item 10.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
D	1,08	0,00	2,89	94,58	1,44	1,29600	-2,55300	0,23340

Fonte: Elaborado pela autora.

Figura D.9: Curva característica e curva de informação do item 10



Fonte: Elaborado pela autora.

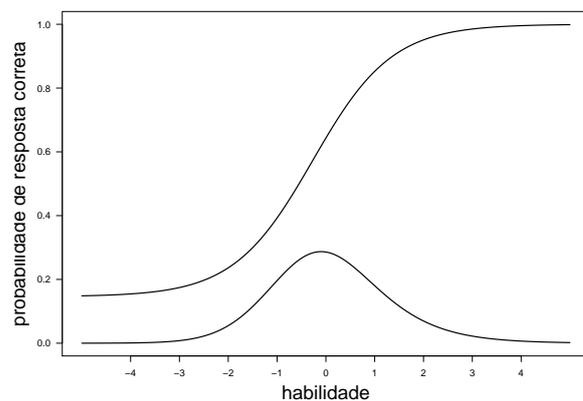
Item 11:

Tabela D.10: Gabarito, percentuais de resposta e parâmetros do item 11.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
B	11,55	62,45	9,03	4,69	11,91	1,23300	-0,26950	0,14600

Fonte: Elaborado pela autora.

Figura D.10: Curva característica e curva de informação do item 11



Fonte: Elaborado pela autora.

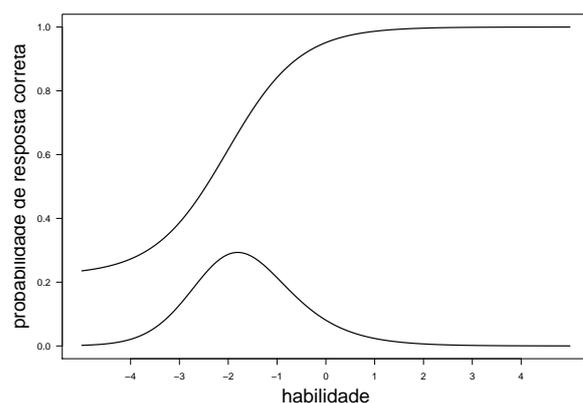
Item 12:

Tabela D.11: Gabarito, percentuais de resposta e parâmetros do item 12.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
E	3,61	0,72	2,53	2,17	90,97	1,33900	-2,02100	0,22130

Fonte: Elaborado pela autora.

Figura D.11: Curva característica e curva de informação do item 12



Fonte: Elaborado pela autora.

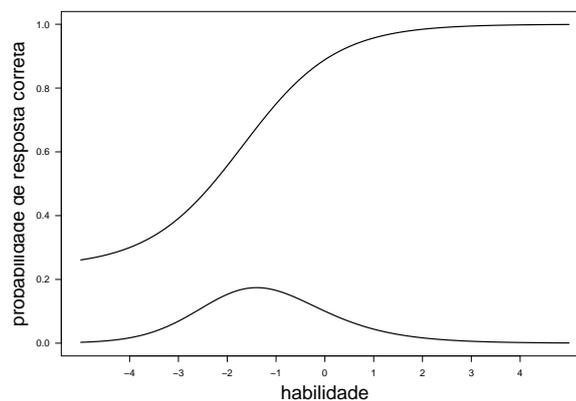
Item 13:

Tabela D.12: Gabarito, percentuais de resposta e parâmetros do item 13.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
D	1,44	0,72	10,47	85,20	2,17	1,04900	-1,68400	0,23830

Fonte: Elaborado pela autora.

Figura D.12: Curva característica e curva de informação do item 13



Fonte: Elaborado pela autora.

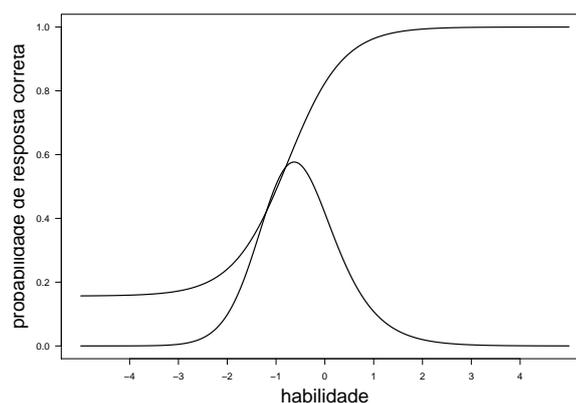
Item 14:

Tabela D.13: Gabarito, percentuais de resposta e parâmetros do item 14.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
A	74,37	9,03	2,17	2,89	11,55	1,76500	-0,75550	0,15650

Fonte: Elaborado pela autora.

Figura D.13: Curva característica e curva de informação do item 14



Fonte: Elaborado pela autora.

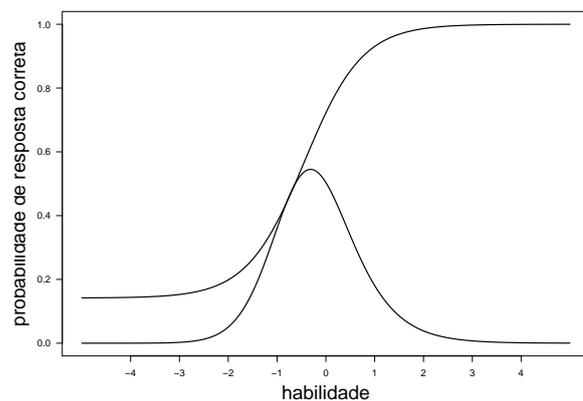
Item 15:

Tabela D.14: Gabarito, percentuais de resposta e parâmetros do item 15.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
E	2,17	7,22	14,80	8,66	66,79	1,69200	-0,43550	0,14170

Fonte: Elaborado pela autora.

Figura D.14: Curva característica e curva de informação do item 15



Fonte: Elaborado pela autora.

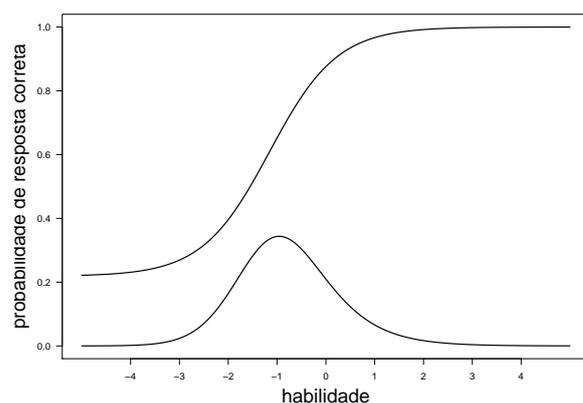
Item 16:

Tabela D.15: Gabarito, percentuais de resposta e parâmetros do item 16.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
E	0,72	3,97	7,94	5,42	81,59	1,44600	-1,15300	0,21860

Fonte: Elaborado pela autora.

Figura D.15: Curva característica e curva de informação do item 16



Fonte: Elaborado pela autora.

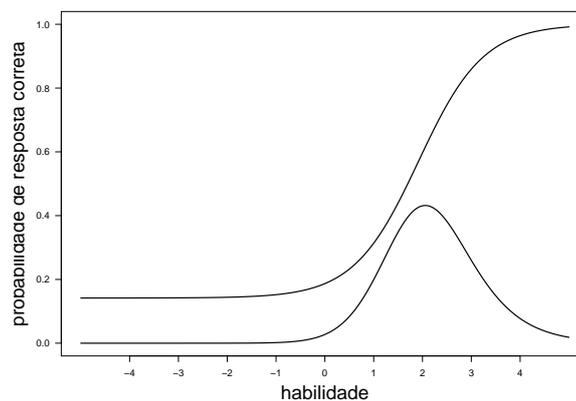
Item 17:

Tabela D.16: Gabarito, percentuais de resposta e parâmetros do item 17.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
C	4,33	37,91	24,91	6,50	24,91	1,50600	1,91900	0,14170

Fonte: Elaborado pela autora.

Figura D.16: Curva característica e curva de informação do item 17



Fonte: Elaborado pela autora.

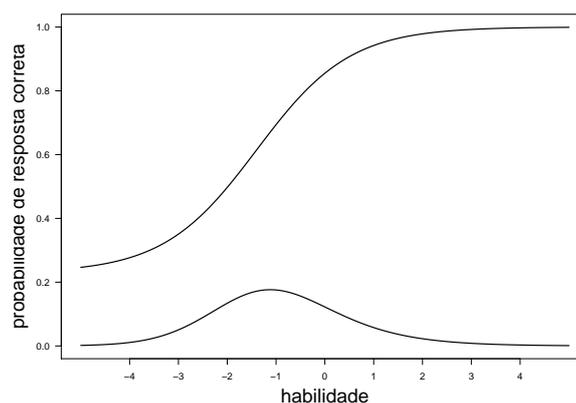
Item 18:

Tabela D.17: Gabarito, percentuais de resposta e parâmetros do item 18.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
E	3,25	1,81	11,55	1,81	81,59	1,04600	-1,39900	0,22900

Fonte: Elaborado pela autora.

Figura D.17: Curva característica e curva de informação do item 18



Fonte: Elaborado pela autora.

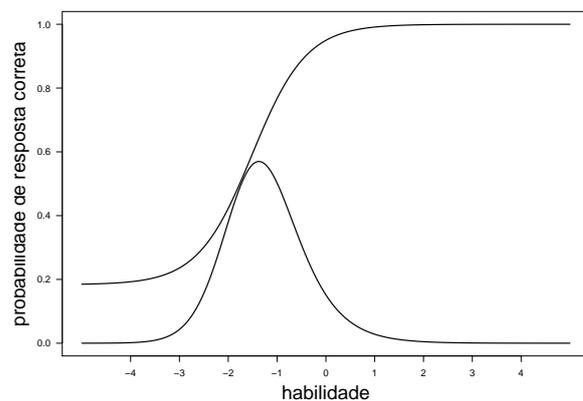
Item 19:

Tabela D.18: Gabarito, percentuais de resposta e parâmetros do item 19.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
C	4,69	3,97	87,73	1,44	2,17	1,80000	-1,51100	0,18350

Fonte: Elaborado pela autora.

Figura D.18: Curva característica e curva de informação do item 19



Fonte: Elaborado pela autora.

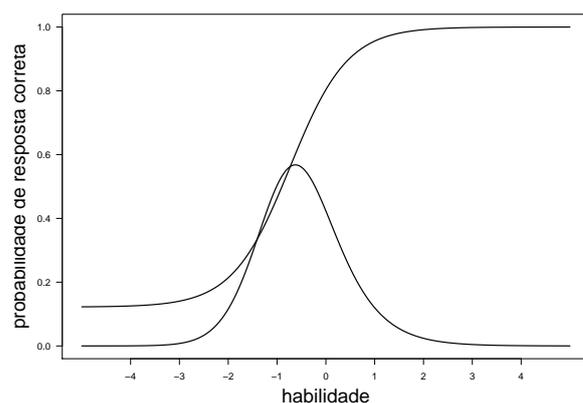
Item 20:

Tabela D.19: Gabarito, percentuais de resposta e parâmetros do item 20.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
D	6,86	13,00	5,78	72,56	1,81	1,69500	-0,73400	0,12200

Fonte: Elaborado pela autora.

Figura D.19: Curva característica e curva de informação do item 20



Fonte: Elaborado pela autora.

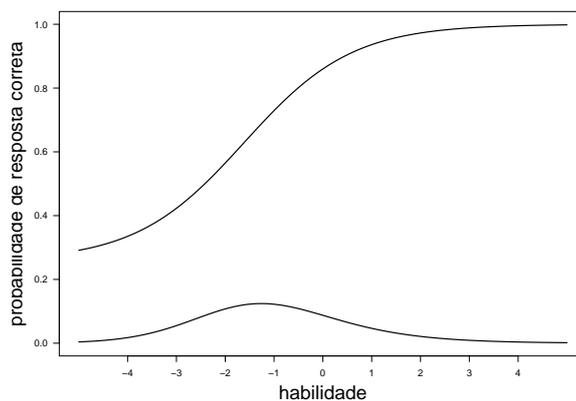
Item 21:

Tabela D.20: Gabarito, percentuais de resposta e parâmetros do item 21.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
C	3,61	0,36	83,03	6,86	6,14	0,90310	-1,61400	0,25800

Fonte: Elaborado pela autora.

Figura D.20: Curva característica e curva de informação do item 21



Fonte: Elaborado pela autora.

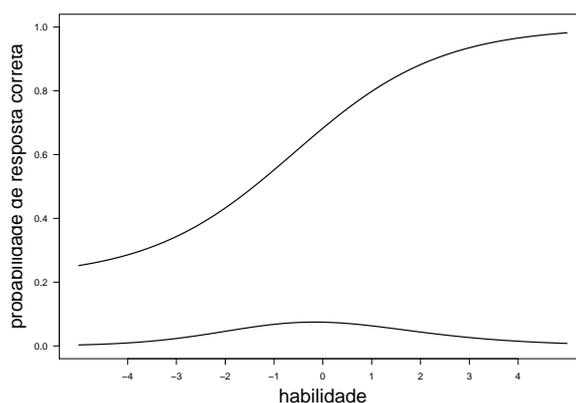
Item 22:

Tabela D.21: Gabarito, percentuais de resposta e parâmetros do item 22.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
B	1,08	67,51	16,61	14,80	0,00	0,67130	-0,58280	0,21350

Fonte: Elaborado pela autora.

Figura D.21: Curva característica e curva de informação do item 22



Fonte: Elaborado pela autora.

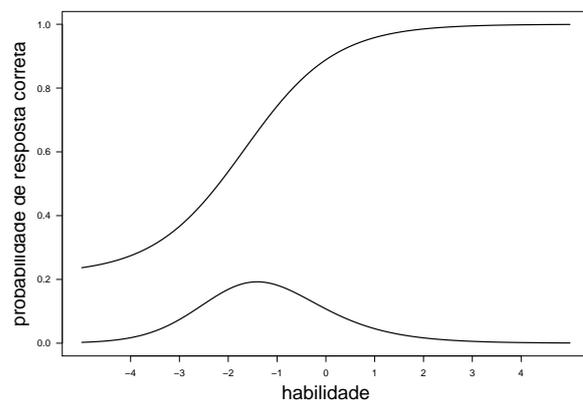
Item 23:

Tabela D.22: Gabarito, percentuais de resposta e parâmetros do item 23.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
C	1,44	5,05	84,84	7,22	1,44	1,07900	-1,66700	0,21580

Fonte: Elaborado pela autora.

Figura D.22: Curva característica e curva de informação do item 23



Fonte: Elaborado pela autora.

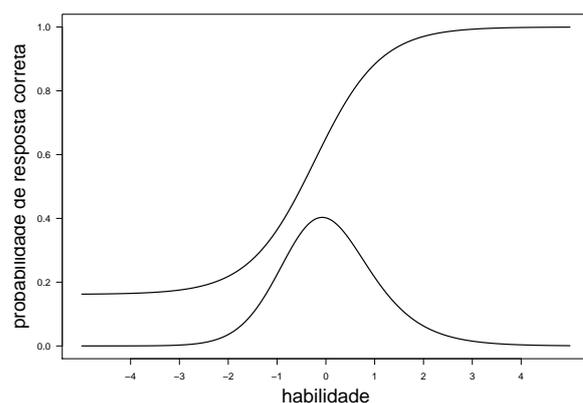
Item 24:

Tabela D.23: Gabarito, percentuais de resposta e parâmetros do item 24.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
B	13,72	62,82	15,88	5,05	2,53	1,48300	-0,22880	0,16170

Fonte: Elaborado pela autora.

Figura D.23: Curva característica e curva de informação do item 24



Fonte: Elaborado pela autora.

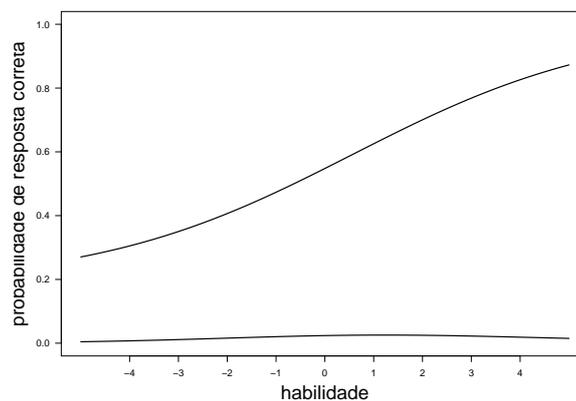
Item 25:

Tabela D.24: Gabarito, percentuais de resposta e parâmetros do item 25.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
D	5,42	3,61	25,27	54,87	10,83	0,38030	0,55930	0,18230

Fonte: Elaborado pela autora.

Figura D.24: Curva característica e curva de informação do item 25



Fonte: Elaborado pela autora.

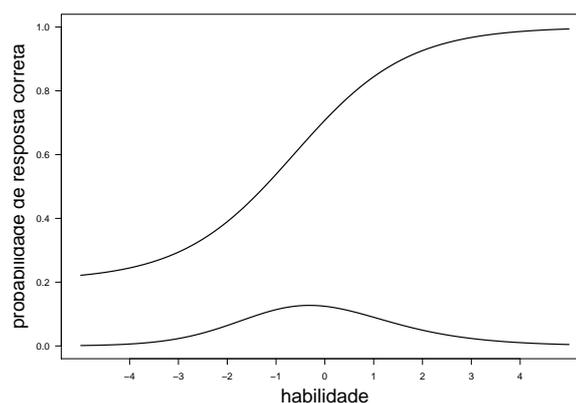
Item 26:

Tabela D.25: Gabarito, percentuais de resposta e parâmetros do item 26.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
C	5,05	14,08	68,95	5,78	6,14	0,86730	-0,62760	0,20400

Fonte: Elaborado pela autora.

Figura D.25: Curva característica e curva de informação do item 26



Fonte: Elaborado pela autora.

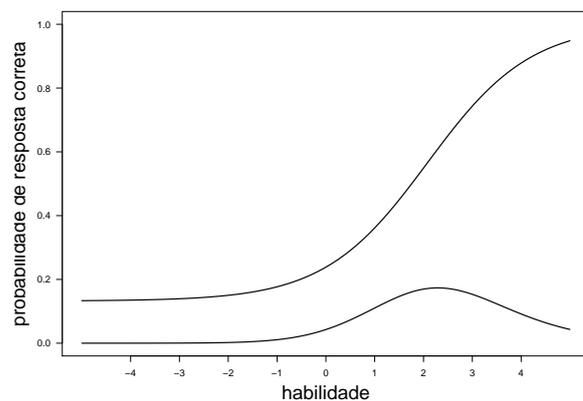
Item 27:

Tabela D.26: Gabarito, percentuais de resposta e parâmetros do item 27.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
A	28,52	28,52	11,55	16,25	14,08	0,94600	2,08300	0,13230

Fonte: Elaborado pela autora.

Figura D.26: Curva característica e curva de informação do item 27



Fonte: Elaborado pela autora.

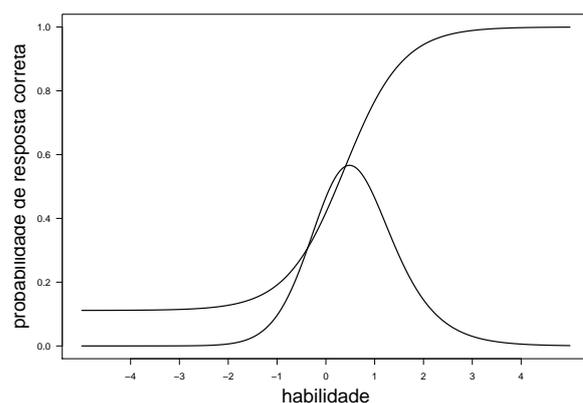
Item 28:

Tabela D.27: Gabarito, percentuais de resposta e parâmetros do item 28.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
D	12,64	18,41	17,33	46,21	3,97	1,67600	0,38140	0,11140

Fonte: Elaborado pela autora.

Figura D.27: Curva característica e curva de informação do item 28



Fonte: Elaborado pela autora.

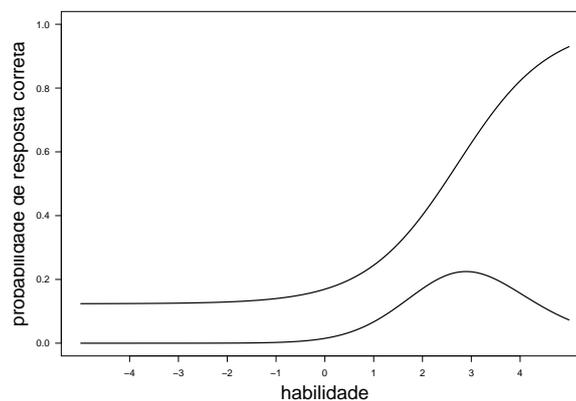
Item 29:

Tabela D.28: Gabarito, percentuais de resposta e parâmetros do item 29.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
B	27,80	20,94	11,55	12,27	27,44	1,06800	2,71700	0,12370

Fonte: Elaborado pela autora.

Figura D.28: Curva característica e curva de informação do item 29



Fonte: Elaborado pela autora.

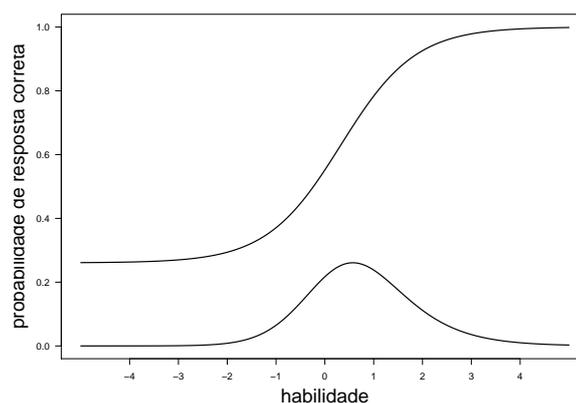
Item 31:

Tabela D.29: Gabarito, percentuais de resposta e parâmetros do item 31.

Indicadores	Percentuais de acertos					Parâmetros TRI		
Gabarito	A	B	C	D	E	a	b	c
A	57,76	23,10	7,58	5,05	6,50	1,31200	0,32820	0,26060

Fonte: Elaborado pela autora.

Figura D.29: Curva característica e curva de informação do item 31



Fonte: Elaborado pela autora.

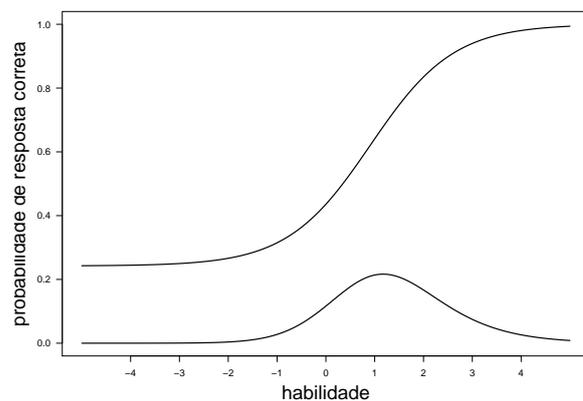
Item 32:

Tabela D.30: Gabarito, percentuais de resposta e parâmetros do item 32.

Indicadores	Percentuais de acertos					Parâmetros TRI		
Gabarito	A	B	C	D	E	a	b	c
B	5,42	48,74	23,47	14,80	6,86	1,17400	0,91170	0,24220

Fonte: Elaborado pela autora.

Figura D.30: Curva característica e curva de informação do item 32



Fonte: Elaborado pela autora.

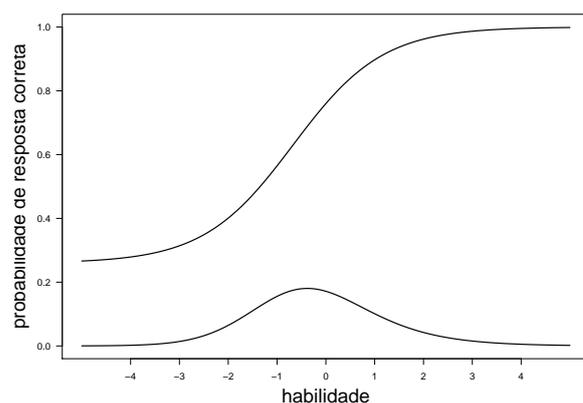
Item 33:

Tabela D.31: Gabarito, percentuais de resposta e parâmetros do item 33.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
D	4,33	12,27	5,42	73,29	4,69	1,09000	-0,67380	0,25970

Fonte: Elaborado pela autora.

Figura D.31: Curva característica e curva de informação do item 33



Fonte: Elaborado pela autora.

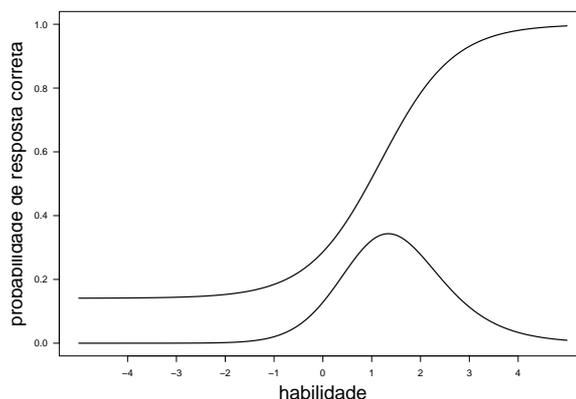
Item 34:

Tabela D.32: Gabarito, percentuais de resposta e parâmetros do item 34.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	Gabarito	A	B	C	D	E	a	b
C	26,35	10,47	35,02	11,55	16,61	1,34200	1,18700	0,14100

Fonte: Elaborado pela autora.

Figura D.32: Curva característica e curva de informação do item 34



Fonte: Elaborado pela autora.

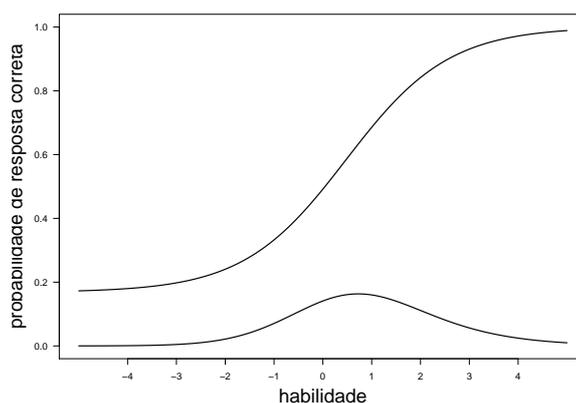
Item 35:

Tabela D.33: Gabarito, percentuais de resposta e parâmetros do item 35.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
A	50,90	29,24	15,16	1,81	2,89	0,94950	0,47590	0,16820

Fonte: Elaborado pela autora.

Figura D.33: Curva característica e curva de informação do item 35



Fonte: Elaborado pela autora.

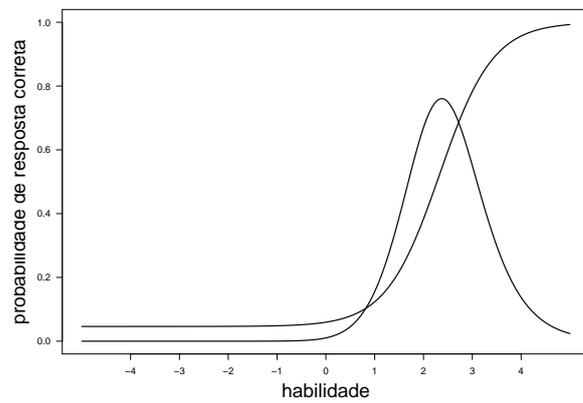
Item 36:

Tabela D.34: Gabarito, percentuais de resposta e parâmetros do item 36.

Indicadores	Percentuais de acertos					Parâmetros TRI		
	A	B	C	D	E	a	b	c
Gabarito								
E	19,13	15,88	43,32	11,19	9,75	1,82500	2,33000	0,04620

Fonte: Elaborado pela autora.

Figura D.34: Curva característica e curva de informação do item 36



Fonte: Elaborado pela autora.

ANEXO

1

Cartão Resposta



Cartão-Resposta

NOME DO CANDIDATO	Nº DE INSCRIÇÃO
ASSINATURA DO CANDIDATO	DOC. DE IDENTIDADE
	DATA

<table border="1"> <thead> <tr> <th colspan="2">IDADE</th> </tr> </thead> <tbody> <tr><td>0 <input checked="" type="checkbox"/></td><td>0 <input checked="" type="checkbox"/></td></tr> <tr><td>1 <input type="checkbox"/></td><td>1 <input type="checkbox"/></td></tr> <tr><td>2 <input type="checkbox"/></td><td>2 <input type="checkbox"/></td></tr> <tr><td>3 <input type="checkbox"/></td><td>3 <input type="checkbox"/></td></tr> <tr><td>4 <input type="checkbox"/></td><td>4 <input type="checkbox"/></td></tr> <tr><td>5 <input type="checkbox"/></td><td>5 <input type="checkbox"/></td></tr> <tr><td>6 <input type="checkbox"/></td><td>6 <input type="checkbox"/></td></tr> <tr><td>7 <input type="checkbox"/></td><td>7 <input type="checkbox"/></td></tr> <tr><td>8 <input type="checkbox"/></td><td>8 <input type="checkbox"/></td></tr> <tr><td>9 <input type="checkbox"/></td><td>9 <input type="checkbox"/></td></tr> </tbody> </table>	IDADE		0 <input checked="" type="checkbox"/>	0 <input checked="" type="checkbox"/>	1 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	5 <input type="checkbox"/>	6 <input type="checkbox"/>	6 <input type="checkbox"/>	7 <input type="checkbox"/>	7 <input type="checkbox"/>	8 <input type="checkbox"/>	8 <input type="checkbox"/>	9 <input type="checkbox"/>	9 <input type="checkbox"/>	<table border="1"> <thead> <tr> <th colspan="2">SEXO</th> </tr> </thead> <tbody> <tr> <td>M <input type="checkbox"/></td> <td>F <input type="checkbox"/></td> </tr> </tbody> </table>	SEXO		M <input type="checkbox"/>	F <input type="checkbox"/>	NÃO ESCREVER NESTA ÁREA
IDADE																												
0 <input checked="" type="checkbox"/>	0 <input checked="" type="checkbox"/>																											
1 <input type="checkbox"/>	1 <input type="checkbox"/>																											
2 <input type="checkbox"/>	2 <input type="checkbox"/>																											
3 <input type="checkbox"/>	3 <input type="checkbox"/>																											
4 <input type="checkbox"/>	4 <input type="checkbox"/>																											
5 <input type="checkbox"/>	5 <input type="checkbox"/>																											
6 <input type="checkbox"/>	6 <input type="checkbox"/>																											
7 <input type="checkbox"/>	7 <input type="checkbox"/>																											
8 <input type="checkbox"/>	8 <input type="checkbox"/>																											
9 <input type="checkbox"/>	9 <input type="checkbox"/>																											
SEXO																												
M <input type="checkbox"/>	F <input type="checkbox"/>																											

Instruções

- Preencha o cartão-resposta com seu nome completo, assinatura e documento de identidade (observe que idade e sexo já fazem parte das respostas).
- Cada questão tem cinco alternativas de resposta: (A), (B), (C), (D) e (E) e apenas uma delas é correta.
- Para cada questão marque a alternativa escolhida no cartão-resposta, preenchendo todo o espaço dentro do retângulo correspondente, com caneta esferográfica azul ou preta.
- Marque apenas uma alternativa para cada questão.
Atenção: questão com mais de uma alternativa marcada será anulada, mesmo que uma das marcações esteja correta.
- Não é permitido o uso de instrumentos de desenho, calculadoras, aparelhos eletrônicos ou quaisquer fontes de consulta.
- Ao final, o cartão-resposta deverá ser entregue juntamente com a prova.

01	<input type="checkbox"/>				
02	<input type="checkbox"/>				
03	<input type="checkbox"/>				
04	<input type="checkbox"/>				
05	<input type="checkbox"/>				
06	<input type="checkbox"/>				
07	<input type="checkbox"/>				
08	<input type="checkbox"/>				
09	<input type="checkbox"/>				
10	<input type="checkbox"/>				
11	<input type="checkbox"/>				
12	<input type="checkbox"/>				
13	<input type="checkbox"/>				
14	<input type="checkbox"/>				
15	<input type="checkbox"/>				
16	<input type="checkbox"/>				
17	<input type="checkbox"/>				
18	<input type="checkbox"/>				
19	<input type="checkbox"/>				
20	<input type="checkbox"/>				
21	<input type="checkbox"/>				
22	<input type="checkbox"/>				
23	<input type="checkbox"/>				
24	<input type="checkbox"/>				
25	<input type="checkbox"/>				
26	<input type="checkbox"/>				
27	<input type="checkbox"/>				
28	<input type="checkbox"/>				
29	<input type="checkbox"/>				
30	<input type="checkbox"/>				
31	<input type="checkbox"/>				
32	<input type="checkbox"/>				
33	<input type="checkbox"/>				
34	<input type="checkbox"/>				
35	<input type="checkbox"/>				
36	<input type="checkbox"/>				
37	<input type="checkbox"/>				
38	<input type="checkbox"/>				
39	<input type="checkbox"/>				
40	<input type="checkbox"/>				

NÃO AMASSE, NÃO DOBRE E NÃO RASURE ESTA FOLHA.

Matriz de Referência

A matriz de referência produzida por Rossi (2015), utilizada neste trabalho, está organizada por temas, sendo estes, espaço e forma, grandezas e medidas e números e operações/álgebra e funções. Cada tema é composto por competências. Cada competência descreve uma habilidade que espera-se que o aluno tenha desenvolvido ao final do ensino médio. Foram consideradas trinta e seis competências relevantes para o acompanhamento das disciplinas de cálculo e similares.

Quadro 2.1: Tema I: Espaço e Forma

Nº	Competência	Origem do Descritor
C1	Identificar a localização de pontos no plano cartesiano	Descritor 6 da Matriz do Ensino Médio do SAEB
C2	Interpretar geometricamente os coeficientes da equação de uma reta	Descritor 7 da Matriz do Ensino Médio do SAEB
C3	Identificar a equação de uma reta apresentada a partir de dois pontos dados ou de um ponto e sua inclinação	Descritor 8 da Matriz do Ensino Médio do SAEB
C4	Relacionar a determinação do ponto de interseção de duas ou mais retas com a resolução de um sistema de equações com duas incógnitas	Descritor 9 da Matriz do Ensino Médio do SAEB
C5	Reconhecer, dentre as equações do segundo grau com duas incógnitas, as que representam circunferências	Descritor 10 da Matriz do Ensino Médio do SAEB

Quadro 2.2: Tema II: Grandezas e Medidas

Nº	Competência	Origem do Descritor
C6	Resolver problemas envolvendo cálculo de perímetro de figuras planas	Descritor 11 da Matriz do Ensino Médio do SAEB
C7	Resolver problemas envolvendo cálculo de área de figuras planas	Descritor 12 da Matriz do Ensino Médio do SAEB
C8	Resolver problemas envolvendo a área total e/ ou volume de um sólido (prisma, pirâmide, cilindro, cone, esfera)	Descritor 13 da Matriz do Ensino Médio do SAEB

Quadro 2.3: Tema III: Números e Operações/ Álgebra e Funções

Nº	Competência	Origem do Descritor
C9	Identificar a localização de números reais na reta numérica	Descritor 14 da Matriz do Ensino Médio do SAEB

(continua)

(continuação)

Nº	Competência	Origem do Descritor
C10	Resolver problemas com números naturais e inteiros envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão e potenciação)	Descritor 19 da Matriz do Ensino Fundamental da Prova Brasil
C11	Identificar a localização de números racionais na reta numérica	Descritor 17 da Matriz do Ensino Fundamental da Prova Brasil
C12	Reconhecer diferentes representações de um número racional	Descritor 21 da Matriz do Ensino Fundamental da Prova Brasil
C13	Identificar fração como representação que pode estar associada a diferentes significados	Descritor 22 da Matriz do Ensino Fundamental da Prova Brasil
C14	Identificar frações equivalentes	Descritor 23 da Matriz do Ensino Fundamental da Prova Brasil
C15	Efetuar cálculos que envolvam operações (adição, subtração, multiplicação, divisão e potenciação) com números racionais	Descritor 25 da Matriz do Ensino Fundamental da Prova Brasil
C16	Resolver problemas com números racionais que envolvam as operações: adição, subtração, multiplicação, divisão e potenciação	Descritor 26 da Matriz do Ensino Fundamental da Prova Brasil
C17	Efetuar cálculos com valores aproximados de radicais	Descritor 27 da Matriz do Ensino Fundamental da Prova Brasil
C18	Calcular o valor numérico de uma expressão algébrica	Descritor 30 da Matriz do Ensino Fundamental da Prova Brasil
C19	Identificar a expressão algébrica que expressa uma regularidade observada em sequências de números ou figuras (padrões)	Descritor 32 da Matriz do Ensino Fundamental da Prova Brasil
C20	Reconhecer expressão algébrica que representa uma função a partir de uma tabela	Descritor 18 da Matriz do Ensino Médio do SAEB
C21	Analisar crescimento/decrescimento, zeros de funções reais apresentadas em gráficos	Descritor 20 da Matriz do Ensino Médio do SAEB
C22	Identificar o gráfico que representa uma situação descrita no texto	Descritor 21 da Matriz do Ensino Médio do SAEB
C23	Identificar e resolver problemas que envolvam funções de primeiro grau	Descritor 19 da Matriz do Ensino Médio do SAEB
C24	Reconhecer a representação algébrica de uma função do primeiro grau, dado o seu gráfico	Descritor 24 da Matriz do Ensino Médio do SAEB
C25	Identificar e resolver problemas de inequações de primeiro grau	Competência acrescentada a Matriz
C26	Identificar e resolver problemas que envolvam equações de segundo grau	Descritor 17 da Matriz do Ensino Médio do SAEB
C27	Identificar e resolver problemas de inequações de segundo grau	Competência acrescentada a Matriz
C28	Resolver problemas que envolvam os pontos de máximo ou de mínimo no gráfico de uma função polinomial de segundo grau	Descritor 25 da Matriz do Ensino Médio do SAEB
C29	Identificar e resolver equações modulares	Competência acrescentada a Matriz
C30	Identificar e resolver inequações modulares	Competência acrescentada a Matriz
C31	Reconhecer o gráfico de uma função polinomial de primeiro grau por meio de seus coeficientes	Descritor 23 da Matriz do Ensino Médio do SAEB
C32	Relacionar as raízes de um polinômio com sua decomposição em fatores do primeiro grau	Descritor 26 da Matriz do Ensino Médio do SAEB

(continua)

(continuação)

Nº	Competência	Origem do Descritor
C33	Identificar a representação gráfica e/ou algébrica de uma função exponencial	Descritor 27 da Matriz do Ensino Médio do SAEB
C34	Identificar a representação gráfica e/ou algébrica de uma função logarítmica, reconhecendo-a como inversa da função exponencial	Descritor 28 da Matriz do Ensino Médio do SAEB
C35	Resolver problemas que envolvam função exponencial	Descritor 29 da Matriz do Ensino Médio do SAEB
C36	Identificar a representação gráfica e/ou algébrica de funções trigonométricas (seno, cosseno, tangente), reconhecendo suas propriedades	Descritor 30 da Matriz do Ensino Médio do SAEB