



UNIVERSIDADE ESTADUAL DO MARANHÃO – UEMA
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
CENTRO DE CIÊNCIAS TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO
MESTRADO EM ENGENHARIA DA COMPUTAÇÃO E SISTEMAS

LIDINALVA DE ALMADA COUTINHO

**FERRAMENTA PARA ESTIMAR A PROFICIÊNCIA DE INDIVÍDUOS
BASEADA NA TEORIA DE RESPOSTA AO ITEM**

São Luís
2013

LIDINALVA DE ALMADA COUTINHO

**FERRAMENTA PARA ESTIMAR A PROFICIÊNCIA DE INDIVÍDUOS
BASEADA NA TEORIA DE RESPOSTA AO ITEM**

Dissertação de Mestrado em Engenharia da Computação e Sistemas, na área de concentração: Informática na Educação, apresentada à Coordenadoria do Programa de Pós-Graduação em Engenharia da Computação e Sistemas da Universidade Estadual do Maranhão, como requisito parcial à obtenção do título de mestra.

Orientador: Prof. Dr. Luís Carlos Costa Fonseca

São Luís
2013

Coutinho, Lidinalva de Almada

Ferramenta para estimar a proficiência de indivíduos baseada na teoria de resposta ao item/Lidinalva de Almada Coutinho – São Luís, 2013.

124 f.

Dissertação (Mestrado) – Curso de Engenharia da Computação e Sistemas, Universidade Estadual do Maranhão, 2013.

Orientador: Prof. Dr. Luís Carlos Costa Fonseca

1.Avaliação. 2.Aprendizagem. 3.Teoria de resposta ao item.
4.Proficiência. 5.Habilidade. Título

CDU: 004:37.091.26

LIDINALVA DE ALMADA COUTINHO

**FERRAMENTA PARA ESTIMAR A PROFICIÊNCIA DE INDIVÍDUOS
BASEADA NA TEORIA DE RESPOSTA AO ITEM**

Dissertação de Mestrado em Engenharia da Computação e Sistemas, na área de concentração: Informática na Educação, apresentada à Coordenadoria do Programa de Pós-Graduação em Engenharia da Computação e Sistemas da Universidade Estadual do Maranhão, como requisito parcial à obtenção do título de mestra.

Aprovada em 31 / 10 / 2013

BANCA EXAMINADORA

Prof. Dr. Luís Carlos Costa Fonseca – UEMA

Orientador

Prof. Dr. Cícero Costa Quarto – UEMA

Examinador Interno

Prof. Dr. Sofiane Labidi – UFMA

Examinador Externo

SÃO LUIS (MA), OUTUBRO DE 2013.

Dedico este trabalho a Deus,
ao meu esposo, Audízio,
às minhas filhas, Vitória e Maria Eduarda,
aos meus pais, José (*in memoriam*) e Rosalvi.

AGRADECIMENTOS

A Deus, pela vida, saúde, sabedoria e oportunidades;

À UEMA, pelas condições oportunizadas e pelo investimento feito para a efetivação da nossa qualificação profissional docente em especial ao CESC/UEMA; com sua infraestrutura e todos os seus servidores, por estar presente em várias etapas da minha formação acadêmica;

Aos professores do PECS/UEMA, por seus ensinamentos;

Aos alunos do Curso Licenciatura em Matemática, por contribuírem para o desenvolvimento da minha pesquisa;

Aos professores do Departamento de Matemática e Física do CESC/UEMA, que me apoiaram e torceram por mim;

Ao meu orientador, Professor Dr. Luís Carlos Costa Fonseca, pela atenção, por acreditar no trabalho e na minha capacidade em desenvolvê-lo, dedicando várias horas a compartilhar sua sabedoria acadêmica e experiência profissional;

Aos membros da banca examinadora por disponibilizarem seus tempos avaliando esta pesquisa e pelas valiosas sugestões e comentários;

A toda minha família e amigos, que estiveram presentes na minha vida, e que contribuíram para minha formação humana e profissional;

Ao meu esposo, Audízio Araújo dos Santos, pelo companheirismo, estímulo e compreensão ao longo do desenvolvimento do trabalho;

Às minhas filhas, Vitória Coutinho dos Santos e Maria Eduarda Coutinho dos Santos, pelo incentivo, amor, inspiração e compreensão durante a minha ausência;

Aos meus pais, José de Almada Coutinho (*in memoriam*) e Rosalvi de Sá Coutinho. Minha eterna gratidão por todos os ensinamentos, por acreditarem no meu potencial, pelo amor e pelos infinitos esforços destinados à minha vida;

Aos meus irmãos, Lindionesa, José, Ronaldo, Rosangela e Marcio, pelo carinho, confiança e amizade;

À minha irmã, Lindonete de Almada Coutinho Silva. A você, meus agradecimentos especiais pela constante disponibilidade, amizade e pelo convívio durante o período do mestrado, demonstrado pelo apoio e amizade;

À minha sobrinha, Carolina de Almada Coutinho Silva, pela alegria, amizade e convívio em todos esses anos;

Ao meu cunhado, Carlos Cesar Silva, pelo apoio, amizade e pelo acolhimento durante o mestrado;

À minha querida tia, Rosilda Coutinho, pelo apoio, força, amizade e pelas sábias palavras de incentivo;

Aos colegas da 1ª Turma do Mestrado, pela parceria nessa jornada de formação profissional;

Ao colega José Wilker Pereira Luz, pelo apoio e parceria nos momentos de estudos;

Ao companheiro Francisco Adelson Ribeiro, pelos momentos de socialização das descontrações, angústias, dúvidas e conflitos nos percalços do curso e da pesquisa;

Ao colega Lanyllo Araujo dos Santos, por ter contribuído de forma significativa para o desenvolvimento desta pesquisa;

Às secretárias Dorinha e Carol, pelo atendimento e pela disposição em ajudar;

À Georgyana Andréia, pela amizade, apoio e parceria;

À Cléia Azevedo, pela amizade, apoio e pelos conhecimentos socializados;

À Shirlane Maria Batista da Silva, pela amizade, companheirismo, incentivo e por está sempre presente nos momentos difíceis;

A todos aqueles que de alguma forma contribuíram para que eu pudesse iniciar, desenvolver e/ou concluir o trabalho.

O que caracteriza a atual revolução tecnológica não é a centralidade de conhecimentos e informação, mas a aplicação desses conhecimentos e dessa informação para a geração de conhecimentos e de dispositivos de processamento/comunicação da informação, em um ciclo de realimentação cumulativo entre a inovação e seu uso.

(Dizard)

RESUMO

Esta dissertação propõe o uso de modelos da Teoria da Resposta ao Item (TRI) no âmbito da avaliação educacional, com o objetivo central de propor uma ferramenta computacional baseada na Teoria de Resposta ao Item que permita verificar a proficiência de indivíduos. Para tanto, aplicou um teste de Matemática, usando como referências as provas do Exame Nacional de Desempenho dos Estudantes (ENADE) 2005/2008/2011. Dessa forma, buscou investigar como avaliar as competências e habilidades dos discentes do curso de Matemática, do Centro de Estudos Superiores de Caxias da Universidade Estadual do Maranhão (CESC/UEMA). Assim, propôs um Sistema, organizado em três módulos, composto por uma ferramenta de autoria de provas, uma ferramenta de aplicação de provas e um corretor de provas, que possibilitam identificar as habilidades dos alunos e fornecer um *feedback* como suporte para a aprendizagem, de acordo com as dificuldades encontradas no teste. Para a análise, utilizou o *software* jMetrik, pelo qual são apresentados o modelo geral para itens dicotomizados, assim como as interpretações dos parâmetros do modelo. Os resultados mostram que a TRI pode ser uma poderosa ferramenta de análise, capaz de gerar informações quantitativas, precisas e robustas.

Palavras-chave: Avaliação. Aprendizagem. Teoria de Resposta ao Item Proficiência. Habilidades.

ABSTRACT

This dissertation proposes the use of models of Item Response Theory (IRT) in the context of educational evaluation, with the main objective to propose a computational tool based on Item Response Theory for verifying the proficiency of subjects. Therefore, it was applied a test of mathematics, using as references the evidence ENADE 2005/2008/2011. Thus, we sought to investigate how to assess the skills and abilities of students of mathematics courses, the CESC / UEMA. Thus, it was proposed a system, organized in three modules, consisting of an authoring tool for evidence, a tool for the application of evidence and broker of evidence that helps identify students' skills and provides a feedback as a support for learning, according to the difficulties encountered in the test. The analysis software, it was used jMetrik, which presents the general model for dichotomized items, as well as the interpretations of the model parameters. The results show that TRI can be a powerful analytical tool capable of generating quantitative information, precise and robust.

Keywords: Evaluation. Learning. Item Response Theory. Proficiency. Skills.

LISTA DE FIGURAS

Figura 1 – Ilustração da medida e seus parâmetros de validade e fidedignidade	34
Figura 2 – Curva Característica do Item com valores especificados	40
Figura 3 – Exemplo de uma Curva Característica do Item (CCI)	43
Figura 4 – Curvas características e de informação de vários itens.....	47
Figura 5 – Captura de tela com listagem das respostas dos alunos às respostas do teste aplicado.....	67
Figura 6 – Detalhamento dos dados do teste	70
Figura 7 – Arquitetura Geral do Sistema	78
Figura 8 – Caso de Uso Autor	78
Figura 9 – Caso de Uso Aluno	79
Figura 10 – Interface para resolução das questões.....	96
Figura 11 – Área do professor na ferramenta.....	97
Figura 12 – Formulário cadastro de provas.....	98
Figura 13 – Formulário de cadastro de questões	99
Figura 14 – Formulário do discente	100
Figura 15 – Interface para resolução das questões.....	101
Figura 16 – Formulário de cadastro de docentes e discentes	102
Figura 17 – Edição de questões.....	103
Figura 18 – Cadastro de provas	104
Figura 19 – Formulário de cadastro de questões	105
Figura 20 – Interface para resolução das questões.....	106

LISTA DE QUADROS E TABELA

Quadro 1 – Visualização panorâmica das perspectivas pedagógicas dos Sistemas AVALWEB, JAVAL e VIRTUAL-TANEB	62
Quadro 2 – Itens analisados	66
Quadro 3 – Representação da frequência das respostas dadas pelos aluno	73
Quadro 4 – Resumo das estatísticas de itens do exame	74
Tabela 1 – Análise dos Itens	107

LISTA DE SIGLAS

ACG – Avaliação dos Cursos de Graduação

AVA – Ambientes Virtuais de Aprendizagem

AVALIES – Avaliação das Instituições de Educação Superior

BM – Banco Mundial

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CCI – Curva Característica do Item

CESC – Centro de Estudos Superiores de Caxias

DIF – Differential Item Functioning

EAD – Ensino a Distância

EM – Maximização de Expectativas

EMV – Estimadores de Máxima Verossimilhança

ENADE – Exame Nacional de Desempenho dos Estudantes

ENCCEJA – Exame Nacional para Certificação de Competências de Jovens e Adultos

ENEM – Exame Nacional dos Estudantes do ensino Médio

FRI – Função de Resposta do Item

GUI – Interface Gráfica do Utilizador

HTML – Linguagem de Marcação de Hipertexto

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais

JPG – Joint Photographic Experts Group

MEC – Ministério da Educação e Cultura

ML1 – Modelo Logístico Unidimensional de um Parâmetro

ML2 – Modelo Logístico Unidimensional de dois Parâmetros

ML3 – Modelo Logístico Unidimensional de três Parâmetros

MRN – Modelo de Resposta Nominal

MYSQL – Linguagem de Consulta Estruturada

NCME – Conselho Nacional de Medição em Educação

OO – Orientação a Objetos

PAPED – Programa de Apoio a Pesquisa em Educação a Distância

PHP – Personal Home Page

PISA – Programa Internacional de Avaliação de Estudantes

PNG – Portable Network Graphics

RENTE – Revista Novas Tecnologias na Educação

SAEB – Sistema Nacional de Educação Básica

SARESP – Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo

SINAES – Sistema Nacional de Avaliação da Educação Superior

SPSS – Statistical Package for the Social Sciences

TCT – Teoria Clássica dos Testes

TRI –Teoria de Resposta ao Item

UEMA – Universidade Estadual do Maranhão

UFRGS – Universidade Federal do Rio Grande do Sul

UML – Linguagem Unificada de Modelagem

VT –VIRTUAL – TANEB

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Objetivos	19
1.1.1	Geral	19
1.1.2	Específicos	19
1.2	Justificativa.....	19
2	AVALIAÇÃO EM LARGA ESCALA	22
2.1	Avaliação no Ensino Superior	25
2.2	Matrizes de Referência: ENADE	27
3	TEORIA DE RESPOSTA AO ITEM – TRI	32
3.1	A Teoria de Resposta ao Item no Brasil	34
3.2	Modelos da Resposta ao Item	35
3.2.1...	Modelos Unidimensionais para Itens Dicotômicos	37
3.2.1.1	ML1 – Modelo logístico (unidimensional) de um parâmetro ou modelo de Rasch.....	37
3.2.1.2	Modelo logístico de 2 parâmetros	38
3.2.1.2.1	Interpretação e representação gráfica.....	41
3.2.1.3	O Modelo Logístico de 3 Parâmetros (ML3)	41
3.2.1.4	Outros modelos para itens dicotômicos	42
3.2.2	Modelos envolvendo um único grupo	43
3.2.3	Modelos para itens não dicotômicos	44
3.2.4	Modelo de Resposta Nominal (Nominal Categories Model)	44
3.3	Função de Informação do Item	45
3.4	Função de Informação do Teste (FIT)	46
3.5	A Escala de Habilidade	48
3.6	Suposições do Modelo: Unidimensional idade e Independência Local	49
3.7	Estimação: uma única população	50
3.7.1	Estimação dos Parâmetros dos Itens	52
3.7.2	Estimação dos Traços Latentes	52
3.8	Equalização	53
3.8.1	Delineamento de grupos não equivalentes com itens comuns	54
3.8.2	Diferentes tipos de equalização	55
3.8.2.1	Um único grupo fazendo uma única prova	55

3.8.2.2	Um único grupo fazendo duas provas totalmente distintas	55
3.8.2.3	Um único grupo fazendo duas provas parcialmente distintas	55
3.8.2.4	Dois grupos fazendo uma única prova.....	56
3.8.2.5	Dois grupos fazendo duas provas totalmente distintas	56
3.8.2.6	Dois grupos fazendo duas provas parcialmente distintas	56
3.9	Diferentes problemas de estimação	56
3.9.1	Quando todos os itens são novos.....	57
3.9.2	Quando todos os itens já estão calibrados	57
3.9.3	Quando alguns itens são novos e outros já estão calibrados	57
3.10	A Escala de Habilidade	58
3.11	Trabalhos Relacionados	59
3.11.1	AVALWEB – Sistema Interativo para Gerência de Questões e Aplicação de Avaliações na Web	59
3.11.2	JAVAL – Ambiente para Avaliação Remota para Ensino à Distância Empregando Agentes Embarcados	60
3.11.3	Modelagem para o Software Virtual-Taneb Baseado na Teoria da Resposta ao Item para Avaliar o Rendimento dos Alunos	60
4	MATERIAL E MÉTODO	64
4.1	Caracterização da Pesquisa	64
4.2	Etapas da Pesquisa	65
4.3	Análise resumida dos itens	73
4.4	Análise da Prova	76
4.5	Interface do aplicador e modelagem do Sistema	76
4.6	Uma proposta de corretor de provas TRI	79
4.6.1	Para entender melhor o problema vamos tomar o seguinte exemplo:	79
5	CONSIDERAÇÕES FINAIS	84
5.1	Contribuições deste trabalho	85
5.2	Trabalhos Futuros	86
	REFERÊNCIAS	87
	APÊNDICES.....	96
	ANEXO.....	113

1 INTRODUÇÃO

Os avanços científicos e tecnológicos têm possibilitado o desenvolvimento de métodos estatísticos mais eficientes no âmbito da pesquisa educacional. Nesse contexto, os sistemas de avaliação de grande escala vêm sendo enriquecidos com técnicas que têm como base a Teoria de Resposta ao Item (TRI), desencadeando verdadeira revolução no setor educacional.

É indiscutível a necessidade de que sejam reavaliados e aperfeiçoados os métodos de avaliação em todos os segmentos de ensino, há uma inquietação evidente nos inúmeros estudos voltados para essa temática, de acordo com: Franco (2001) e Gatti (1996). As reflexões acerca dessa temática procuram desvelar novos modelos de propostas de avaliação, que se colocam como perspectivas para a otimização do processo de ensino e de aprendizagem, com vistas à aprendizagem mais significativa.

Considerando o exposto, o curso Licenciatura em Matemática do Centro de Estudos Superiores de Caxias da Universidade Estadual do Maranhão (CESC/UEMA) vem passando por grandes dificuldades, tanto no âmbito do quadro de docentes efetivos (não atende à demanda), quanto no das questões estruturais relacionadas aos laboratórios. Refletir e discutir sobre a formação de professores de Matemática tem se tornado cada vez mais frequente, visto que nas três últimas avaliações do Exame Nacional de Desempenho de Estudantes (ENADE) (2005, 2008, 2011), de acordo com relatórios do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o Curso tem apresentado resultados insatisfatórios, sendo motivo de grande preocupação para toda a comunidade acadêmica do referido Centro. Para tanto, o presente estudo fundamenta-se nos seguintes teóricos: Andrade et al. (2000), Creswell (2007), Dias Sobrinho (2004), Franco (2001), Gil (2010), Libâneo (1991), Meyer (2010), Pasquali (2003), Perrenoud (1999), Richardson (2010), Vianna (2005), dentre outros.

Nessa perspectiva, como o ENADE é uma realidade sobre a qual os alunos, em sua maioria, desconhecem os métodos utilizados no processo avaliativo, causa uma inquietação que poderá ser minimizada caso esses alunos tenham a oportunidade de vivenciar algo semelhante. Para tanto, é necessário considerar a experiência vivida por alunos e professores, valorizando a cultura na qual eles estão inseridos, e partindo desses conhecimentos e experiências, elaborar um conjunto de

estratégias que atenda e crie condições de potencializar os conhecimentos existentes, favorecendo seu bom desempenho e, conseqüentemente, melhorando a situação do curso Licenciatura em Matemática junto ao Ministério da Educação e Cultura (MEC) e a comunidade acadêmica do CESC/UEMA, buscando articular a cultura existente no cotidiano de cada um com a cultura erudita. E uma das formas que o aluno tem para ser incitado a desenvolver e incentivar a sua autonomia é conhecer ferramentas que busquem continuamente novos caminhos e sejam capazes de criar e recriar conhecimentos significativos.

Para tanto, propõe-se a seguinte questão norteadora: Como avaliar as competências e as habilidades dos discentes do curso de Matemática, do CESC/UEMA, utilizando uma ferramenta computacional que identifique as habilidades dos alunos e forneça um *feedback* como suporte para a aprendizagem, de acordo com as dificuldades encontradas no teste, empregando técnicas da Teoria de Resposta ao Item (TRI)?

Compreende-se que a discussão acerca da avaliação de competências e habilidades, requer uma compreensão das peculiaridades do processo de ensino e de aprendizagem como um todo. É notório que avaliar por si só, não garantirá ao aluno o acesso e o domínio do conhecimento científico, nem tão pouco habilidades necessárias para desempenhar suas funções a contento. Todavia, uma ferramenta avaliativa que utiliza técnica da Teoria de Resposta ao Item, possibilitará uma maior articulação entre os diversos saberes disciplinares e o uso e aplicação desses saberes de forma contextualizada.

Destarte, o diferencial para a realização desta pesquisa, é a necessidade de recursos computacionais eficientes, para aferir o nível de conhecimento dos alunos e fornecer um *feedback* que auxiliará o discente na busca e compreensão de conhecimentos matemáticos disponibilizados durante a avaliação. Outro ponto relevante é a possibilidade do aluno aprender de acordo com o seu próprio ritmo de aprendizagem, além de ajudá-lo a desenvolver atividades de aprofundamento, contribuindo para a aprendizagem de novas habilidades, conhecimentos e valores.

1.1 Objetivos

1.1.1 Geral

Propor uma ferramenta computacional que permita a automação da avaliação do nível de conhecimento dos alunos do ensino superior, segundo os moldes exigidos pelo ENADE.

1.1.2 Específicos

Como principais objetivos, podem-se listar os seguintes:

- i) Apresentar os fundamentos básicos da Teoria da Resposta ao Item, bem como, suas vantagens e oportunidades dentro no processo de avaliação educacional;
- ii) Propor a modelagem de *software* que trabalhe baseado nos conceitos da Teoria de Resposta ao Item na avaliação da aprendizagem;
- iii) Desenvolver uma ferramenta de avaliação de aprendizagem que utilize a Teoria da Resposta ao Item e que abstraia a complexidade de sua utilização.

1.2 Justificativa

A Avaliação Educacional é um elemento imprescindível no processo de ensino e de aprendizagem. Entretanto, o processo educativo é complexo e as diferentes concepções que se tem de avaliação são discutidas continuamente e dependendo de como se compreende e se internaliza essas concepções, avalia-se o sujeito e desconsidera-se, conseqüentemente, na maioria das vezes, a relação que deve existir entre a ação e reflexão, tendo em vista, suas complexidades e rotinas, sua vivência e, principalmente, sobre questões que se destina a construção do conhecimento.

Assim, é cada vez mais crescente a aplicação de técnicas de avaliação utilizadas nas mais variadas áreas do conhecimento, principalmente na avaliação educacional, na qual vem se destacando o uso da Teoria de Resposta ao Item, que

trata de modelos para os traços latentes, ou seja, características do sujeito que não podem ser observadas diretamente. Esse tipo de variável deve ser inferida a partir da análise de variáveis secundárias que estejam relacionadas a ela. O mais interessante nessa são as formas de representar a relação entre a probabilidade de um indivíduo dar uma resposta a um item e seus traços latentes, proficiências ou habilidades na área de conhecimento avaliada.

Nas duas últimas décadas, a TRI vem sendo cada vez mais utilizada graças aos resultados satisfatórios que se tem obtido em vários países do mundo, tornando uma ferramenta predominante no campo de testes estatísticos. No Brasil, a TRI foi introduzida em 1995, na análise dos dados do Sistema Nacional de Ensino Básico (SAEB). Através da TRI, foi possível comparar e colocar os desempenhos de alunos de 3ª, 4ª e 8ª séries do ensino em uma escala única de conhecimento.

A partir dos resultados obtidos no SAEB, outras avaliações em larga escala, como as avaliações internas, representadas pelas provas usadas pelos docentes como forma de verificar o desempenho acadêmico dos alunos; e as avaliações externas, representadas pelas avaliações institucionais, tais como o Exame Nacional do Desempenho dos Estudantes (Prova), tendo como uma das finalidades verificar as proficiências básicas dos concluintes dos cursos de graduações, o Exame Nacional do Ensino Médio (ENEM) e outros exames de avaliação estaduais como Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP) e Ambientes Virtuais de Aprendizagem do Paraná (AVA) têm fornecido informações valiosas às políticas educacionais, dentre outros.

Considerando o exposto, esta pesquisa justifica-se pela necessidade do desenvolvimento de uma ferramenta computacional de avaliação, utilizando recursos da TRI que contemplem um modelo pedagógico que atendam aos anseios da comunidade estudantil do curso de Matemática do CESC/UEMA, no sentido de familiarizá-los com processo de avaliação do ENADE.

Com base nessa proposição, o presente estudo está estruturado em cinco capítulos. **O primeiro capítulo** contempla a parte introdutória do trabalho, na qual se abordam a justificativa, objetivos da pesquisa e a proposta de estruturação da dissertação.

O segundo capítulo, intitulado “Avaliação em Larga Escala”, em que são descritos os caminhos trilhados na pesquisa, caracterizando tanto a avaliação no

contexto institucional brasileiro, quanto a sua importância no processo de ensino e de aprendizagem.

O terceiro capítulo, “Teoria de Resposta ao Item”, apresenta os aportes conceituais da pesquisa, enfatizando os diferentes modelos e suas respectivas aplicações. Ressalta-se, ainda, que nesse contexto, focaliza-se a perspectiva histórica da TRI.

O quarto capítulo, denominado “Material e Método”, remete aos procedimentos metodológicos que serão desenvolvidos, incluindo a proposta da arquitetura do sistema de avaliação, com a descrição geral do sistema e a modelagem dos requisitos funcionais e não funcionais através dos casos de uso. Nesse capítulo são destacadas: a análise dos dados produzidos na pesquisa, a proposta do corretor de provas, e registram-se, também, através de uma leitura crítico-analítica, os achados da pesquisa. De modo similar, realiza-se a análise interpretativa das representações, a partir de reflexões acerca dos resultados encontrados.

O quinto capítulo trata das considerações finais, apresenta-se a conclusão da pesquisa, apontando algumas reflexões e constatações do estudo e sugestão para trabalhos futuros.

2 AVALIAÇÃO EM LARGA ESCALA

Desde a década de 1990, o processo de avaliação educacional vem se destacando, a partir de duas dimensões: uma interna, avaliação da aprendizagem realizada pelo professor como parte da sua prática pedagógica; e a outra externa, que é caracterizada pela avaliação do desempenho escolar, em larga escala, realizada por agente externo à escola. Para Wiebusch (2012), são necessárias as discussões no contexto escolar, abordando questões referentes tanto à avaliação interna como à externa, para que de fato possam cumprir com o seu papel, na obtenção de resultados efetivos de melhoria da aprendizagem dos alunos através de iniciativas que permitam a implementação de ações mais condizentes com a oferta de uma educação de qualidade e promoção da equidade de oportunidades educacionais.

Nesse sentido, de acordo com Dias Sobrinho (2004), a avaliação é assumida como um elemento que traz implícito o compromisso social com a garantia de padrões de excelência na formação e na produção de conhecimentos, garantindo a todos que possam se beneficiar do ensino superior; com o desenvolvimento de consciência crítica e com a melhoria das condições de vida da sociedade.

Portanto, é fundamental considerar a avaliação interna e a externa como alternativa na reflexão da prática educativa e a necessidade de informar os resultados para a sociedade. Penin (2009) discorre sobre o assunto, declarando que:

No âmbito interno, possibilita a avaliação como instrumento de ação formativa, levando instituições e os professores a refletirem a respeito de suas práticas e de seus objetivos e, assim, a melhorar sua ação docente e sua identidade profissional. Por outro, em âmbito externo, oferece informações para que tanto os pais quanto a sociedade, especialmente os sistemas de ensino, possam efetivar um relacionamento produtivo com a instituição escolar. Apurar os usos da avaliação, comparar resultados e comportamento de entrada dos alunos em cada situação e contexto social e institucional é da maior importância para não homogeneizar processos que são de fato diferentes. (PENIN, 2009, p. 23-24).

No contexto brasileiro, Castro (2009) afirma que tais tendências influenciou na criação do ENADE, do SAEB, do ENEM, da Prova Brasil, além da participação no Programa Internacional de Avaliação de Estudantes (PISA).

Dessa forma, as informações obtidas a partir de um sistema de avaliação têm papel fundamental sobre os rumos do sistema de ensino, além de procurar garantir a fidedignidade das informações oferecidas, é importante garantir a reflexão sobre esses resultados e, conseqüentemente, melhoria na sua produção, considerando os atores participantes do processo, o aprimoramento de métodos, instrumentos e logística de realização da avaliação. Outro fato que merece destaque são os avanços em relação aos aspectos metodológicos e institucionais, no acompanhamento de políticas educacionais e a associação entre avaliação e promoção de políticas de equidade. De acordo com Libâneo (1991):

[...] avaliação como uma componente do processo de ensino que visa, através da verificação e qualificação dos resultados obtidos, a determinar a correspondência destes com os objetivos propostos e, daí, orientar a tomada de decisões em relação às atividades didáticas seguintes. (LIBÂNEO, 1991, p. 196).

O processo de avaliação corresponde a um conjunto de informações sobre determinada realidade, e está presente no cotidiano escolar. Para tanto são utilizados instrumentos desde uma simples observação até as mais sofisticadas ferramentas que permitem avaliar as competências e habilidades esperadas, em diversos níveis de complexidade.

Confere-se que na realização de uma avaliação é importante considerar: o que vai ser avaliado; estabelecer os critérios e as condições para a avaliação; selecionar as técnicas e instrumentos de avaliação; realizar a aferição dos resultados. Além disso, pode-se recorrer aos mais variados métodos de avaliação, que serão selecionados de acordo com o tipo de avaliação, podendo ser: diagnóstica, somativa ou formativa. Segundo Ramos (1999), um bom instrumento de avaliação é aquele que informa bem, é discriminador, é consistente, não é arbitrário. Uma boa metodologia de avaliação deve ir muito além do medir, do verificar, do classificar, precisa oferecer muito mais do que um índice quantitativo, deve permitir compreender, conhecer, interpretar, identificar, situar, entender o que e como se aprendeu.

Desse modo, é importante considerar além das características familiares dos alunos, para o sucesso ou fracasso escolar, os fatores escolares. Ignorá-los implica em enfraquecer a possibilidade de investimentos e intervenção governamentais que sejam capazes de assegurar permanência dos discentes com sucesso escolar.

A avaliação em larga escala no Brasil, bem como seus resultados não implicou em avaliar a qualidade do ensino, posto que o rendimento dos alunos é apenas um, não menos importante, dos elementos constitutivos dos processos educativos composto por variáveis complexas como vivência/experiência na/cultura escolar, a socialização, a sociabilidade, as interações, trocas. Araújo e Fernandes (2009, p. 138).

Assim sendo, o SAEB avalia a qualidade, a equidade e a eficiência do ensino e da aprendizagem, no âmbito do Ensino Fundamental e do Ensino Médio. Aplicado a cada dois anos, o SAEB também avalia o que os alunos sabem e são capazes de fazer em diversas situações de seu período escolar, levando em conta as condições existentes nas escolas brasileiras. Para isso, Bonamino e Franco (1999) enfatizam que são utilizados instrumentos específicos como provas aplicadas a alunos de escolas selecionadas por amostragem em todas as unidades da Federação, nas quais é medido o desempenho acadêmico dos estudantes, através de questionários pelos quais são investigados os fatores intra e extraescolares associados ao desempenho dos alunos. Assim, o SAEB é um instrumento essencial de apoio a todos que lidam com a educação em nosso país.

Na política educacional brasileira. Bonamino e Franco (1999, p. 15) esclarecem que:

Na definição dos objetivos gerais do SAEB pode-se perceber que a formulação mais descentralizada e participativa adotada no primeiro Ciclo, que enfatiza os aspectos processuais ligados ao desenvolvimento de uma cultura de avaliação no interior do sistema educacional, foi sendo secundarizada em favor de uma definição mais centralizada no âmbito do MEC.

Verifica-se que a principal finalidade se constitui em conhecer e construir parâmetros de rendimento dos estudantes, com base em propostas curriculares já existentes. O que gerou muitas críticas, havendo necessidade de modificações no segundo ciclo do SAEB.

Notadamente, as avaliações criadas nos últimos anos pelo Ministério da Educação e Cultura (MEC), como o ENADE, conhecido até o ano de 2003 como “provão”, foram colocadas para a sociedade com a finalidade de melhorar a qualidade dos cursos superiores oferecidos aos estudantes e também verificar as proficiências básicas dos concluintes dos cursos de graduação.

2.1 Avaliação no Ensino Superior

A avaliação, na educação superior está além do aplicar testes, levantar medidas, selecionar e classificar alunos. É importante considerar a dimensão política e emancipatória com vistas a uma proposta contemporânea que faça frente à complexidade das demandas da sociedade. Portanto, “Avaliar, para muitos de nós, professores da educação superior, é uma das atividades pedagógicas mais difíceis de realizar [...]” (ABREU; MASETTO, 1990, p. 108).

A avaliação universitária compreende um estudo teórico sobre revisão de conceitos, entendimentos e experiências, além de um apanhado da situação institucional a respeito de políticas, programas, propostas e realizações existentes, dadas as características de cada área de ensino/conhecimento. Para contemplar uma proposta que na perspectiva de Dias Sobrinho (2000) atenda as exigências da comunidade acadêmica foi criado um sistema de avaliação, com a finalidade de integrar e articular, de forma coerente, concepções, objetivos, metodologias, práticas, agentes da comunidade acadêmica e de instâncias do governo.

Nesse cenário, o Sistema Nacional de Avaliação da Educação Superior (SINAES), instituído pela Lei 10.861 de 14 de abril de 2004, tem por objetivo “assegurar o processo nacional de avaliação das instituições de educação superior, dos cursos de graduação e do desempenho acadêmico de seus estudantes, nos termos do artigo 9º, VI, VIII e IX, da Lei nº 9.394, de 20 de dezembro de 1996”.

Dessa forma, o INEP é o responsável pela avaliação do ensino superior no Brasil. No âmbito do SINAES, o INEP realiza um conjunto de avaliações integrando três modalidades principais de instrumentos avaliativos, aplicados em diferentes momentos:

Avaliação das Instituições de Educação Superior (AVALIES) – Centro de referência e articulação do sistema de avaliação, momento no qual nas Instituições de ensino e a autonomia avaliativa terão seu maior alcance, e assim identificarão e trabalharão suas singularidades e princípios norteadores;

Avaliação dos Cursos de Graduação (ACG) – MEC, avalia os cursos de graduação por meio de instrumentos e de procedimentos que incluem visitas *in loco* de comissões externas;

Avaliação do Desempenho dos Estudantes (ENADE) – Avaliação dos estudantes, que é realizada pela aferição do desempenho acadêmico obtido nos

cursos de graduação, junto a estudantes do final do primeiro e do último ano do curso. Abreu Júnior (2009, p. 267) considera que:

Se universidade é uma instituição dinâmica, seus processos avaliativos também o são. O respeito às singularidades institucionais não recomenda a adoção de medidas ou procedimentos universais, incontestáveis ou absolutos. Se a avaliação permite uma aprendizagem institucional, sendo ela mesma um desses processos, deve conduzir o seu desenvolvimento pela própria avaliação, ou seja, ela deve prever a sua meta – avaliação, em uma busca da coerência interna com seus princípios e finalidades. Trata-se de um conceito que aponta a superação de uma avaliação meramente descritiva da universidade, em favor de um processo de identificação e hierarquização dos problemas, que possibilite o envolvimento dos atores também na busca das soluções e que, por tanto, colabore para o desenvolvimento humano da instituição e a qualidade acadêmica de seus processos.

O ENADE foi instituído em 2004, como um novo modelo de exame aplicado aos estudantes de graduação. Como um dos instrumentos de avaliação que compõem o SINAES, ele tem como objetivo aferir:

O desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares do respectivo curso de graduação, suas habilidades para ajustamento às exigências decorrentes da evolução do conhecimento e suas competências para compreender temas exteriores ao âmbito específico de sua profissão ligado à realidade brasileira e mundial e a outras áreas do conhecimento. (BRASIL, 2004).

O exame é realizado obedecendo a um ciclo avaliativo, com duração de três anos, entre os cursos avaliados. É importante enfatizar que em um processo avaliativo democrático, os sujeitos envolvidos devem ter consciência sobre qual é o ideal, e entender a serviço de quem está esse ideal. Avaliar é um processo dinâmico de reflexão sobre o que fazemos. É um movimento constante e permanente entre ação e reflexão. Desse modo, avaliação e aprendizagem estão intrinsecamente entrelaçados. Uma boa avaliação deve começar, a partir do diagnóstico da realidade. A avaliação qualitativa ultrapassa a avaliação quantitativa, porém uma não pode dispensar a outra.

Em uma proposta de avaliação é importante considerar as práticas dos professores e vivências dos alunos, como ressalta Dias Sobrinho (2000):

Uma prática democrática que se vá construindo coletivamente e esteja sempre orientada à produção da qualidade educativa, ao melhoramento institucional e à permanente formação dos agentes da comunidade. Não uma proposta fechada e que reivindique a pretensão da aprovação geral e incontestada, e sim esforço de construção de consensos relativos sobre princípios básicos que permitam certos graus de possibilidade de trabalho coletivo. (DIAS SOBRINHO, 2000, p. 17).

Entretanto, a compreensão da realidade pelos sujeitos envolvidos deve ser compartilhada. Tornando-se um veículo de informação e de investigação. Uma boa ferramenta de avaliação é aquela que informa bem, a ideia da avaliação qualitativa vai além do que foi proposto inicialmente. De acordo com Demo (2004, p. 156),

A avaliação qualitativa pretende ultrapassar a avaliação quantitativa, sem dispensar esta. Entende que no espaço educativo os processos são mais relevantes que os produtos, não fazendo jus à realidade, se reduzida apenas às manifestações empiricamente mensuráveis. [...] A avaliação qualitativa gostaria de chegar até à face qualitativa da realidade, ou pelo menos de se aproximar dela.

A competência para transformar a avaliação em uma prática democrática, requer novas relações, ressaltando assim, a necessidade de repensar a prática pedagógica, propiciando situações de aprendizagem focadas em ferramentas que poderão oportunizar um novo sentido à avaliação escolar, através da interação, colaboração e cooperação entre os pares, em uma perspectiva colaborativa.

2.2 Matrizes de Referência: ENADE

Para a realização de avaliações externas em larga escala foram elaboradas Matrizes de Referência, que descrevem o que se pretende avaliar com os itens contidos nas provas realizadas pelos alunos. As Matrizes referem-se aos conteúdos considerados referências, nos quais é contemplada a noção de competência, sendo tomada como princípio de organização curricular, engendrando habilidades que os alunos devem dispor e demonstrar nas avaliações.

Tais competências são entendidas como modalidades estruturais da inteligência, ou melhor, como o conjunto de operações e ações mentais que uma pessoa utiliza para estabelecer relações com os objetos e entre as situações, fenômenos, etc. Perrenoud (1999, p. 7), define competência como “[...] uma

capacidade de agir eficazmente em um determinado tipo de situação, apoiada em conhecimentos, mas sem limitar-se a eles”.

As matrizes de referência diferenciam na quantidade de competências e na descrição da habilidade avaliada. De acordo com os documentos oficiais, dentre eles as Diretrizes Curriculares Nacionais, há a necessidade de centrar o ensino e a aprendizagem no desenvolvimento de competências e habilidades por parte do aluno, em detrimento da centralização conceitual e conteudista.

A conceituação de competência e habilidades assumida pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) pode ser entendida por:

Competências são as modalidades estruturais da inteligência, ou melhor, ações e operações que utilizamos para estabelecer relações com e entre objetos, situações, fenômenos e pessoas que desejamos conhecer. As habilidades decorrem das competências adquiridas e referem-se ao plano imediato do “saber fazer”. Através das ações e operações, as habilidades aperfeiçoam-se e articulam-se, possibilitando nova reorganização das competências. (INEP, 1999, p. 9).

Portanto, a competência representa o modo como fazemos convergir nossas necessidades e articulamos nossas habilidades em favor de um objeto ou solução de um problema, expresso num desafio.

As Matrizes de Referência para a Avaliação apresentam competência, a partir da definição de Perrenoud (1999), como sendo a faculdade de mobilizar um determinado conjunto de recursos cognitivos, tais como os saberes, as habilidades e as informações, para solucionar com pertinência e eficácia uma série de situações.

Para fazer uma avaliação em larga escala, ou seja, elaborar as questões/itens, que constarão na prova, é necessário que se tenha uma matriz de referência, em que estarão explicitadas as habilidades a serem avaliadas, consideradas fundamentais, possíveis de serem avaliadas em um teste de múltipla escolha.

As Matrizes de Referência são compostas por um conjunto de descritores, que contemplam dois pontos básicos do que se pretende avaliar:

- i) o conteúdo programático e;
- ii) o nível de operação mental necessário para a habilidade avaliada.

Tais descritores são selecionados para compor a matriz, considerando-se aquilo que pode ser avaliado por meio de itens de múltipla escolha. O descritor é o detalhamento de uma habilidade cognitiva, e este considera o grau de complexidade do item, que está sempre associado a um conteúdo que o estudante deve dominar na etapa de ensino em análise. Esses descritores são expressos da forma mais detalhada possível, para a mensuração por meio de aspectos que podem ser observados.

Os descritores são responsáveis pela articulação entre conteúdos, competências e habilidades. Acredita-se que a análise dos resultados do ENADE poderá auxiliar os cursos para a melhoria do processo educacional, por meio da reflexão e discussão do trabalho realizado, avaliando o que está bem, o que precisa ser melhorado e o que está faltando em termos de competências e habilidades. Compreende-se, que os resultados das avaliações externas são uma excelente oportunidade de análise para a educação (re)pensar o seu fazer pedagógico e, no coletivo, construir ações, metas, realizar intervenções pedagógicas para elevar o desempenho dos alunos, promovendo, assim, a equidade e a melhoria da qualidade da educação, bem como a necessidade de a instituição conhecer com clareza o perfil dos alunos, identificando, compreendendo a realidade em que se encontram e os rumos que almejam atingir.

A utilização concreta e objetiva dos resultados precisa ser o foco da equipe gestora, dos professores e de todos os integrantes da instituição para uma tomada de decisão sobre a vida escolar dos alunos, principalmente na dimensão pedagógica, pois não basta só diagnosticar, devemos usar os resultados para uma reflexão coletiva, para o redirecionamento e o planejamento de ações que auxiliem os professores no cotidiano da sala de aula. Tais reflexões fazem pensar que é preciso encontrar novos caminhos para explorar efetivamente os resultados das avaliações externas a favor da aprendizagem, cumprindo, assim, os objetivos a que essas avaliações se propõem. O uso dos resultados das avaliações externas pela instituição deve colaborar para repensar todos os aspectos e gerar transformações.

Vianna (2005, p. 17) reforça: Os resultados das avaliações não devem ser usados única e exclusivamente para traduzir certo desempenho escolar. A sua utilização implica em servir de forma positiva na definição de novas políticas públicas, de projetos de implantação e modificação de currículos, de programas de formação continuada dos docentes e, de maneira decisiva, na definição de

elementos para a tomada de decisões que visem a provocar um impacto, ou seja, mudanças no pensar e no agir dos integrantes do sistema.

De acordo com Sampaio (2010, p. 63), “Esquecemos que somos uma unidade que envolve mente, sentimentos e corpo físico”. A Educação para a integridade do Ser é uma proposta educacional preocupada em educar na e para a autoformação; não apenas em relação aos aspectos racionais (cognitivos) ou à construção do conhecimento, mas que essa construção seja voltada para a autoformação nas dimensões que integram nossa condição humana, de modo especial os aspectos relativos ao corpo, às emoções e à espiritualidade, vários estudos científicos comprovam que essas dimensões são desprestigiadas.

No que diz respeito à importância da avaliação interna e da externa, como alternativa para refletir sobre a prática educativa e a necessidade de informar os resultados para todos, Penin (2009, p. 23-24) reforça que:

No âmbito interno, possibilita a avaliação como instrumento de ação formativa, levando instituições e os professores a refletirem a respeito de suas práticas e de seus objetivos e, assim, a melhorar sua ação docente e sua identidade profissional.

Por outro, em âmbito externo, a avaliação oferece informações para que tanto os pais quanto a sociedade, especialmente os sistemas de ensino, possam efetivar um relacionamento produtivo com a instituição escolar. Apurar os usos da avaliação, comparar resultados e comportamento de entrada dos alunos em cada situação e contexto social e institucional é da maior importância para não homogeneizar processos que são de fato diferentes.

De modo geral, as avaliações educacionais fazem uso de dois tipos de instrumentos: testes e questionários. Os questionários são os mais utilizados, pois, por meio deles, adquirem-se informações que permitem, entre outros aspectos, traçar um perfil das condições escolares. Enquanto que os testes têm a função de oferecer informações acerca das habilidades cognitivas construídas pelos alunos. No cenário atual, torna-se necessário e imprescindível repensar e redimensionar os processos avaliativos, proporcionando ao sistema educacional conhecimentos e ações compatíveis com a nova demanda das tendências educacionais.

Allen e Yen, (2002), teóricos da educação, preocupam-se em conceber, desenvolver e aplicar métodos avaliativos que representem de forma mais fiel

possível os resultados da aprendizagem. Nesse contexto, a teoria da medida tem tido implicações diretas no delineamento, interpretação e resultados de pesquisas e avaliações educacionais. Para Pasquali (2003), tal teoria desenvolve uma discussão epistemológica em torno da utilização dos números no estudo científico dos fenômenos naturais.

Produzir um instrumento para medir uma variável nas ciências sociais não é uma tarefa fácil e inclui uma série de etapas que devem ser seguidas rigorosamente. Segundo Andrade et al. (2010), deve-se considerar as seguintes etapas:

- i) conceituação dos comportamentos que definem operacionalmente o construto em questão;
- ii) elaboração de itens que acessem o construto;
- iii) administração dos itens elaborados para amostras pré-definidas;
- iv) refinamento do instrumento baseado em análises dos itens;
- v) realização de estudos de validade e confiabilidade.

É importante considerar essas etapas para assegurar que os escores em um instrumento sejam consistentes e realmente acessem o construto que se pretende avaliar. No campo da medição, há predomínio de duas teorias, a saber: a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). Esses conceitos serão mais detalhados na seção seguinte.

3 TEORIA DE RESPOSTA AO ITEM (TRI)

Teoria de Resposta ao Item (TRI), segundo Andrade et al (2000), é um conjunto de modelos matemáticos que tem por finalidade representar a probabilidade de um indivíduo dar uma certa resposta a um item em um determinado teste, levando em consideração os parâmetros do item e a habilidade do respondente. Para tanto, essa função se apresenta de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item.

Considerando o exposto, destaca-se que a TRI surgiu a partir de discussões teóricas sobre a viabilidade de se comparar as habilidades e os conhecimentos de indivíduos submetidos a diferentes tipos de prova. A ferramenta estatística utilizada à época para realizar a referida análise era a Teoria Clássica dos Testes (TCT) que, segundo Pasquali (2003), tem como principal objetivo analisar a resposta final, ou seja, o resultado do que os itens dizem sobre o sujeito. Tal interpretação não estava correspondendo de forma satisfatória aos resultados apresentados.

A partir de discussões sobre a fidedignidade apresentada pela TCT relacionada a conceitos de equidade, foram realizados vários estudos, dentre os quais se destacam os trabalhos de Lord (1952), que defendia as vantagens teóricas dos modelos da TRI em procedimentos que estabelecem a equivalência dos resultados de provas, ou seja, considerava partes e probabilidades que geram o resultado final.

Após algumas aplicações desse modelo, Lord (1952), sentiu a necessidade da incorporação de um parâmetro que tratasse do problema do acerto casual. Assim, surgiu o modelo de três parâmetros. Anos mais tarde, Birnbaum (1968) substituiu, em ambos os modelos, a função ogiva normal pela função logística, matematicamente mais conveniente, pois é uma função explícita dos parâmetros do item e de habilidade e não envolve integração. Independentemente do trabalho de Lord e Rasch (1960), propôs o modelo unidimensional de um parâmetro, expresso também como modelo de ogiva normal e, também, mais tarde, descrito por um modelo logístico por Wright (1968).

Samegima (1969) propôs o modelo de resposta gradual com o objetivo de obter mais informação das respostas dos indivíduos do que simplesmente se eles deram respostas corretas ou incorretas aos itens. Bock (1972), Andrich (1978),

Masters (1982) e Muraki (1992) também propuseram modelos para mais de duas categorias de resposta, assumindo diferentes estruturas entre essas categorias.

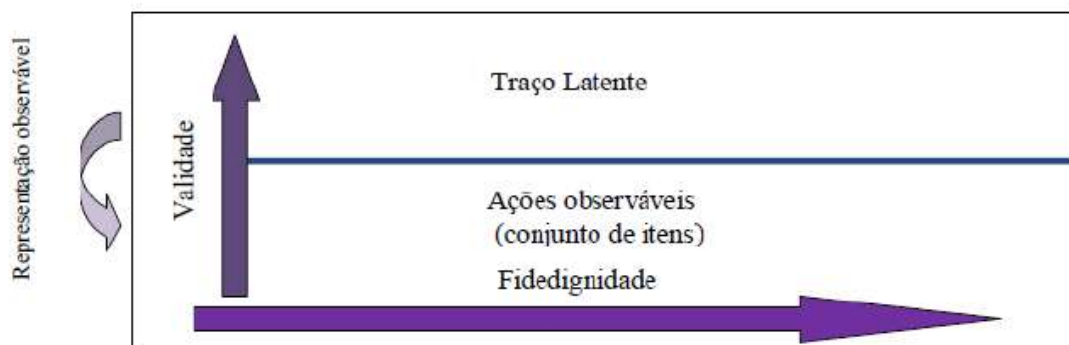
Entretanto, Bock e Zimowski (1997) introduziram os modelos logísticos de um, dois e três parâmetros para duas ou mais populações de respondentes. Em se tratando de avaliação educacional, é importante observar as características comuns a uma determinada população, tais características que podem variar, dependendo dos objetivos do estudo e, portanto, podem ou não ser relevantes para a diferenciação de populações. A introdução desses modelos trouxe novas possibilidades para as comparações de rendimentos de duas ou mais populações submetidas a diferentes testes com itens comuns, conforme discutido em Hedges e Vevea (1997) e Andrade (1999), por exemplo.

Dessa forma, verifica-se que um ponto crítico na TRI é a estimação dos parâmetros envolvidos nos modelos, em particular quando se necessita estimar tanto os parâmetros dos itens quanto as habilidades. Inicialmente, a estimação era feita através do método da máxima verossimilhança conjunta que envolve um número muito grande de parâmetros a serem estimados simultaneamente e, conseqüentemente, grandes problemas computacionais.

Compreende-se, assim, que esses problemas se direcionam para a necessidade da realização de novas pesquisas que se destacaram em 1970, quando Bock e Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas. Na primeira etapa estimam-se os parâmetros dos itens, assumindo-se certa distribuição para as habilidades. Na segunda etapa, assumindo os parâmetros dos itens conhecidos, estimam-se as habilidades. Apesar do avanço que esse método trouxe para o problema, havia uma exigência que todos os parâmetros dos itens fossem estimados simultaneamente.

A partir das ideias mencionadas, visualiza-se, de acordo com Rodrigues (2012), por meio da Figura 1 a seguir, a ilustração de como a medida e seus parâmetros de validade e fidedignidade necessitariam ser idealizados, observando os traços latentes por meio de uma representação observável.

Figura 1 – Ilustração da medida e seus parâmetros de validade e fidedignidade



Fonte: Adaptada de Pasquali (1997, p. 83).

Para garantir a validade e fidedignidade de uma escala de medida, essa deve ser construída com respaldo teórico e prático, obedecendo algumas etapas que são fundamentais para o desenvolvimento e compreensão do Sistema a ser considerado. O desempenho (habilidade) dos respondentes é conhecido através dos resultados das respostas dadas à avaliação em questão.

3.1 A Teoria de Resposta ao Item no Brasil

A TRI vem sendo utilizada em larga escala desde o seu surgimento, com inúmeras aplicações em diferentes áreas do conhecimento, porém somente na década de 1990 deu-se início às primeiras aplicações no Brasil. De acordo com Andrade et al (2000), através da pesquisa AVEJU, da Secretaria de Estado da Educação de São Paulo, e continuaram no Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP), no Sistema de Avaliação da Educação Básica (SAEB) do INEP/MEC em 1995 e, posteriormente, foi implementado também no Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA), Prova Brasil e ENEM.

Nessas aplicações, a TRI tem mostrado a sua potencialidade no que diz respeito à avaliação educacional, por meio da construção de uma escala comparável, permitindo o acompanhamento do progresso do conhecimento adquirido pelo aluno ao longo do tempo, como tem sido feito nos países do Primeiro Mundo (ANDRADE; TAVARES; VALLE, 2000).

Embora tenha começado na área de avaliação educacional, a TRI se expandiu rapidamente para as mais diversas áreas que têm como objetivo a avaliação, principalmente quando há necessidade de evidenciar a comparação do desempenho entre duas ou mais populações de indivíduos. No Brasil, tem-se buscado obter o mesmo sucesso que a TRI teve e tem nos países do primeiro mundo.

A TRI trata o problema da estimação da habilidade e conhecimento de um examinando de forma essencialmente diferente: o enfoque das análises desvincula-se das provas (como na TCT) e concentra-se nos Itens; se na TCT as estatísticas dos itens dependem da população dos examinandos e da prova à qual os itens pertencem, na TRI, de acordo com Valle (1999), cria-se o conceito de que os parâmetros dos itens, obtidos no processo estatístico de "calibração" dos parâmetros de dificuldade, discriminação e acerto casual dos itens são características próprias dos mesmos. Costuma-se considerar que a característica de medição dos Itens, representados por seus parâmetros, são invariantes no tempo com algumas ressalvas.

Uma característica fundamental para a viabilidade de comparação da habilidade e conhecimento de examinandos submetidos a provas diferentes é que a Teoria da Resposta ao Item modela a probabilidade de acerto a um item, também conhecida como Curva Característica do Item (CCI), através de uma função não linear do conhecimento dos examinandos. Essa característica da modelagem da Teoria da Resposta ao Item é de grande importância, portanto, é possível comparar o conhecimento dos examinandos submetidos a provas diferentes sendo necessário apenas que as provas meçam as mesmas características; essa propriedade é essencialmente útil para sistemas de avaliações nos quais é possível submeter uma grande quantidade de tópicos de uma matéria em sala de aula, oportunizando uma melhor visualização sobre o ensino de vários tópicos, com os alunos respondendo apenas a um conjunto pequeno dos itens utilizados na avaliação, evitando tornar as provas muito extensas.

3.2 Modelos da Resposta ao Item

Na TRI é fundamental considerar o traço latente, variável que deve ser inferida a partir da observação de variáveis secundárias que estejam relacionadas a

ela, variáveis não observáveis ou habilidades. Por exemplo, a inteligência, a habilidade em executar uma tarefa, ansiedade, o nível de entendimento de texto, entre outras características inerentes ao ser humano, elas não podem ser medidas diretamente como o peso ou altura de uma pessoa.

Portanto, a TRI considera a base de todo seu desenvolvimento dois postulados:

- i) a performance de um respondente em um teste pode ser prevista por um conjunto de fatores inerentes ao indivíduo, chamados de habilidade ou traço latente;
- ii) a relação entre a habilidade do indivíduo e a probabilidade de escolha no item pode ser escrita por uma função característica ou Curva Característica Item (CCI), parametrizada por características do item de um teste (1 – poder de discriminação; 2 – poder de dificuldade).

De acordo com Andrade et al. (2000), entre os modelos propostos na literatura, considera-se fundamentalmente três fatores:

- i) a natureza do item — dicotômicos ou não dicotômicos;
- ii) o número de populações envolvidas — apenas uma ou mais de uma;
- iii) e a quantidade de traços latentes que está sendo medida — apenas um (unidimensionalidade) ou mais de um (multidimensionalidade).

Na educação, os modelos de TRI mais utilizados são os modelos logísticos para itens dicotômicos, que podem ser utilizados tanto para a análise de itens de múltipla escolha dicotomizados, quanto para a análise de itens abertos, quando avaliados também de forma dicotomizada, e se diferenciam pelo número de parâmetros utilizados para descrever o item: um, dois e três parâmetros, que consideram respectivamente, **a dificuldade do item, a dificuldade e a discriminação, e a dificuldade, discriminação e probabilidade de acerto casual ou “chute”** (ANDRADE et al., 2000). Será apresentado com maior detalhe o modelo logístico unidimensional de três parâmetros, que consideram, respectivamente:

- i) somente a dificuldade do item;

- ii) a dificuldade e a discriminação;
- iii) a dificuldade, a discriminação e a probabilidade de resposta correta dada por indivíduos de baixa habilidade. Haja vista que os outros dois modelos são obtidos facilmente a partir desse que é mais complexo.

Além da funcionalidade que o modelo logístico de 3 parâmetros apresenta, favorece uma visão mais ampla da situação a ser analisada, uma vez que é na visão de Andrade et al. (2000), é considerado o mais completo e portanto os outros dois podem ser facilmente obtidos a partir dele.

3.2.1 Modelos Unidimensionais para Itens Dicotômicos

Os primeiros modelos de resposta ao item foram propostos na década de 1950. Esses modelos consideravam que uma única habilidade de um dado grupo estava sendo medida por um teste no qual os itens eram dicotômicos ou dicotomizados (corrigidos como certo ou errado). Os modelos logísticos para itens dicotômicos são os modelos de resposta ao item mais utilizados, sendo que há basicamente três tipos, que se diferenciam pelo número de parâmetros que utilizam para descrever o item. Eles são conhecidos como os modelos logísticos de um, dois e três parâmetros, que são apresentados a seguir:

3.2.1.1 ML1 – Modelo logístico (unidimensional) de um parâmetro ou modelo de Rasch.

Segundo García (1999), Rasch começou seu trabalho em medidas educacionais e psicométricas por volta de 1940. Por volta da década de 1950, Rasch, usando a função de Poisson, desenvolveu dois modelos, um para leitura de testes e um modelo para aproveitamento e inteligência de testes, com a finalidade de produção de scores.

Com base em seus trabalhos, a motivação de Rasch foi representar a probabilidade de resposta como função da habilidade do respondente e a característica do item. Seja θ o parâmetro de habilidade do respondente j e b a dificuldade do item i .

O sucesso do respondente é a razão entre sua habilidade e a soma da habilidade com a dificuldade do item. Dessa forma, Rasch construiu:

$$P_i(U_{ij} = 1/\theta) = \frac{\frac{\theta_j}{\delta_i}}{\frac{\theta_j}{1-\delta_i}} = \frac{\theta_j}{\theta_j + \delta_i} \quad (1)$$

Em que será utilizado:

$$\theta_j = \theta \quad \text{e} \quad \delta_i = b_i$$

Tomando em uma escala logarítmica para os parâmetros, que será adotada por convenção daqui por diante. Tem-se o modelo unidimensional de um parâmetro:

$$P(U_{ij} = 1/\theta_j) = \frac{1}{1 + e^{-(\theta_j - b_i)}} \quad (2)$$

No qual:

- i) U_{ij} – é uma variável dicotômica que assume os valores 1, quando o indivíduo j responde corretamente o item i, ou 0 quando o indivíduo j não responde corretamente ao item i;
- ii) $P(U_{ij} = \frac{1}{\theta_j})$ – é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i. O parâmetro de dificuldade (ou de posição) do item i, medido na mesma escala da habilidade, é denotado por b_i .

Desse modo, conclui-se que sob o ponto de vista de Guewehr (2007), se além de não existir resposta ao acaso e todos os itens tiverem o mesmo poder de discriminação, tem-se o modelo logístico de um parâmetro, também conhecido como Modelo de Rasch.

3.2.1.2 Modelo logístico de dois parâmetros (ML2)

O modelo proposto por Birnbaum (1968), a partir da substituição da função de distribuição normal, proposta no modelo de Lord (1952), pela função logística, de acordo com Silva (2013), pressupõe a relação monótona entre o valor da variável latente (a ser estimada) do indivíduo e a sua probabilidade de escolha por uma das duas alternativas, segundo uma função de distribuição logística parametrizada por coeficientes que representam determinadas características do item. É muito natural no contexto de avaliação educacional, cuja a variável latente é identificada com a habilidade cognitiva do aluno, e as possibilidades de escolha são acertar ou não o item.

Assim, admita que U_{ij} seja uma variável aleatória dicotômica assumindo os valores 0 ou 1. No caso específico de um teste educacional o valor 0 está associado a uma resposta errada e, o valor 1 a uma resposta certa por parte do aluno. O modelo de dois parâmetros expressa a relação entre a variável latente θ e a resposta dada ao item da seguinte forma:

$$P\left(U_{ij} = \frac{1}{\theta_j}\right) = \frac{1}{1+e^{-Da_i(\theta_j-b_i)}} \quad (3)$$

Em que:

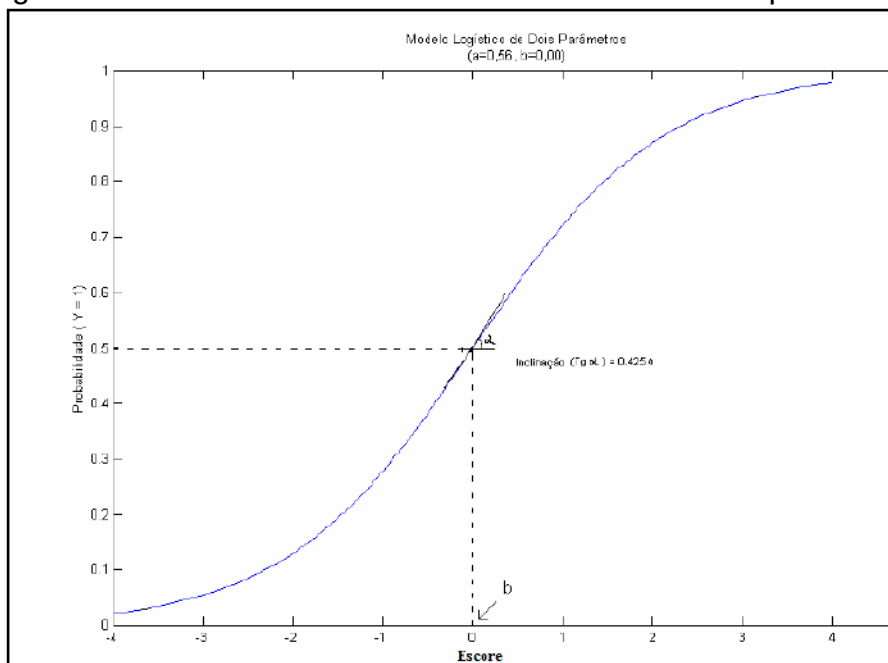
- i) a_i – é o parâmetro de discriminação (ou de inclinação) do item i, com valor proporcional à inclinação da Curva Característica do Item – CCI no ponto b_i ;
- ii) D – é um fator de escala constante e igual a 1. Utiliza-se o valor 1,7 quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal, diz-se, então, que o modelo está na métrica normal;
- iii) i – representa o número do item;
- iv) j – o respondente.

Como se pode notar, o parâmetro b_i representa o ponto na escala de habilidade em que um examinando tem 50% de probabilidade de responder ao item i corretamente. Num contexto mais geral, b_i representa o valor da variável latente θ , para o qual há 50% de chance de escolha da resposta representada por $U_{ij} = 1$ pelo

indivíduo. É fácil observar que se (6) for derivada em relação a θ , a função resultante atinge seu máximo em $\theta = b_i$ com um valor diretamente proporcional a a_i ($0.425 a_i$), indicando que a inclinação da curva do modelo atinge seu maior valor em que a probabilidade de ocorrer uma resposta representada por $U_{ij} = 1$ (isto é, a resposta correta, no caso de modelos para avaliação educacional) é 0.5.

Portanto, quanto maior for o valor do parâmetro a_i , mais sensível torna-se o modelo a variações na habilidade em torno de seu ponto de dificuldade. Isto é, maiores valores para o parâmetro a_i produzirão maior capacidade de distinção entre dois indivíduos com habilidades diferentes no nível da escala em torno do nível de dificuldade do item. Por isso, ele é conhecido como parâmetro de discriminação do item. A seguir, apresenta-se a denominada curva característica de um item, isto é, a representação dos valores, sob a forma de gráfico, de um particular modelo θ , enfatizando as propriedades de seus parâmetros:

Figura 2 – Curva Característica do Item com valores especificados



Fonte: Pereira (2004)

Muito embora a motivação principal na proposição desse modelo tenha sido o de sua utilização em avaliação educacional, Pereira (2004), faz algumas considerações sobre possibilidades de uso. É o caso de, por exemplo, utilizá-lo para a construção de um índice que mede a condição socioeconômica de indivíduos de

uma população. Podem ser consideradas como variáveis indicadoras da condição socioeconômica, a posse de determinados bens como, por exemplo, eletrodomésticos, automóvel, etc.

3.2.1.2.1 Interpretação e representação gráfica

Note que $P(U_{ij} = 1/\theta_j)$ pode ser vista como a proporção de respostas corretas ao item i dentre todos os indivíduos da população com habilidade θ_j . A relação existente entre $P(U_{ij} = 1/\theta_j)$ e os parâmetros do modelo é mostrada na Figura 2, que é chamada de Curva Característica do Item (CCI).

De acordo com o modelo proposto, baseia-se no fato de que indivíduos com maior habilidade possuem maior probabilidade de acertar o item e que essa relação não é linear. De fato, pode-se perceber a partir do gráfico acima que a CCI tem forma de “S” com inclinação e deslocamento na escala de habilidade definidos pelos parâmetros do item.

A escala da habilidade é uma escala arbitrária na qual o importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude. O parâmetro b é medido na mesma unidade da habilidade e o parâmetro c não depende da escala, pois se trata de uma probabilidade, e como tal, assume sempre valores entre 0 e 1.

Na realidade, o parâmetro b representa a habilidade necessária para uma probabilidade de acerto igual a $(1 + c)/2$. Assim, quanto maior o valor de b , mais difícil é o item, e vice-versa.

3.2.1.3 Modelo logístico de três parâmetros (ML3)

Dos modelos propostos pela TRI, o modelo logístico unidimensional de três parâmetros (ML3) é atualmente o mais utilizado e é dado por:

$$P\left(U_{ij} = \frac{1}{\theta_j}\right) = c_i + (1 - c_i) \frac{1}{1 + e^{-D_{ai}(\theta_j - b_i)}}, \quad (4)$$

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

- i) U_{ij} – é uma variável dicotômica que assume os valores 1, quando o indivíduo j responde corretamente o item i , ou 0 quando o indivíduo j não responde corretamente ao item i . θ_j representa a habilidade (traço latente) do j -ésimo indivíduo;
- ii) $P(U_{ij} = 1/\theta_j)$ – é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i e é chamada de Função de Resposta do Item (FRI);
- iii) b_i – é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala da habilidade;
- iv) a_i – é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item no ponto b_i ;
- v) c_i – é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (muitas vezes referido como a probabilidade de acerto casual);
- vi) D – é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal.

3.2.1.4 Outros modelos para itens dicotômicos

Dois outros modelos podem ser facilmente obtidos a partir do modelo logístico de três parâmetros. Por exemplo, quando não existe possibilidade de acerto ao acaso, pode-se considerar $c = 0$ no modelo anterior e tem-se o chamado modelo logístico unidimensional de dois parâmetros (ML2), dado por:

$$P(U_{ij} = 1/\theta_j) = \frac{1}{1 + e^{-D a_i(\theta_j - b_i)}} \quad (5)$$

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

Se além de não existir resposta ao acaso, ainda tivermos todos os itens com o mesmo poder de discriminação, tem-se o chamado modelo logístico

unidimensional de 1 parâmetro (ML1), também conhecido como modelo de Rasch. Este modelo é dado por:

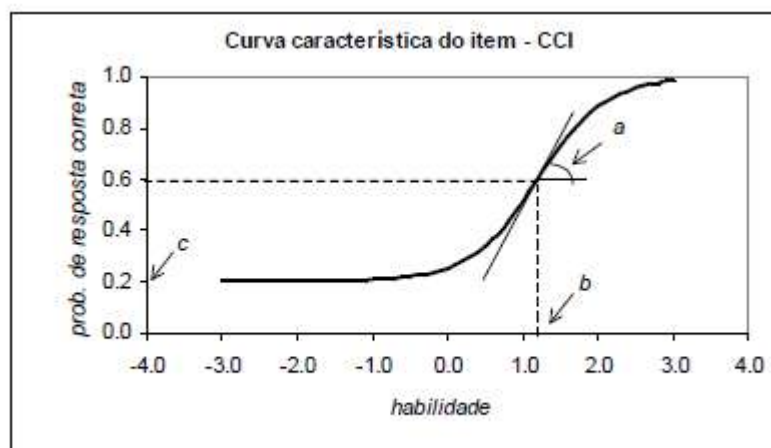
$$\left(U_{ij} = \frac{1}{\theta_j} \right) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}, \quad (6)$$

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

3.2.2 Modelos envolvendo um único grupo

O parâmetro c representa a probabilidade de um aluno com baixa habilidade responder corretamente o item e é muitas vezes referido como a probabilidade de acerto ao acaso.

Figura 3 – Exemplo de uma Curva Característica do Item (CCI)



Fonte: Andrade et al. (2000)

Então, quando não é permitido “chutar”, c é igual a 0 e b representa o ponto na escala da habilidade cuja a probabilidade de acertar o item é 0,5. O parâmetro a é proporcional à derivada da tangente da curva no ponto de inflexão. Assim, itens com a negativo não são esperados sob esse modelo, uma vez que indicariam que a probabilidade de responder corretamente o item diminui com o aumento da habilidade.

No entanto, os baixos valores de a indicam que o item tem pouco poder de discriminação (alunos com habilidades bastante diferentes têm aproximadamente a mesma probabilidade de responder corretamente ao item) e os valores muito altos

indicam itens com curvas características muito “íngremes”, que discriminam os alunos basicamente em dois grupos: os que possuem habilidades abaixo do valor do parâmetro b e os que possuem habilidades acima do valor do parâmetro b .

3.2.3 Modelos para itens não dicotômicos

Aqui são incluídos os modelos tanto para a análise de itens abertos (de resposta livre) quanto para a análise de itens de múltipla escolha que são avaliados de forma graduada, ou seja, itens que são elaborados ou corrigidos de modo a ter-se uma ou mais categorias intermediárias, ordenadas entre as categorias certo e errado. Nesse tipo de item não se considera somente se o indivíduo respondeu à alternativa correta ou não, mas também se leva em conta qual foi a resposta dada por ele.

3.2.4 Modelo de Resposta Nominal (Nominal Categories Model)

Com o objetivo de obter estimativas mais precisas dos traços latentes de indivíduos submetidos a testes de múltipla escolha. Bock (1972) propôs o Modelo de Resposta Nominal (MRN), que define a probabilidade de resposta para cada uma das alternativas de um item em função do traço latente e dos parâmetros dos itens, ao contrário dos modelos dicotômicos que consideram apenas se o indivíduo respondeu ou não corretamente o item.

Para Bock (1972), através da alternativa incorreta é possível melhorar as estimativas dos traços latentes. No MRN a presença de um grande número de parâmetros torna a estrutura matemática complexa. E, comparando outros modelos para itens poliatômicos, ele não estabelece nenhuma ordem de grandeza entre as alternativas do item.

Bock (1972), De Ayala e Sava-Bolesta (1999) mostraram que o aumento na precisão das estimativas dos traços latentes pode ser obtido quando as informações sobre o construto contidas nas categorias incorretas são incluídas no modelo.

Nesse contexto, a utilização das categorias incorretas no processo de estimação pode aumentar a precisão das estimativas dos traços latentes e fornecer informação sobre o nível de compreensão do indivíduo. Levine e Drasgow (1983) mostraram que existe uma relação entre a escolha de uma categoria e os níveis do

traço latente do indivíduo, ou seja, a forma de escolher determinada categoria está diretamente relacionada ao traço latente do indivíduo.

Pinheiro et al. (2010) apresentaram o MRN como uma alternativa para a parametrização dos itens de múltipla escolha, comparando o ajuste do MRN com os MRDs, além de realizar a análise das questões, dos distratores e da precisão dos dados dicotômicos e politômicos.

Segundo De Ayala e Sava-Bolesta (1999), a maioria dos trabalhos que utilizam os modelos da TRI para analisar testes de múltipla escolha são baseados em modelos dicotômicos, tais como o modelo logístico de dois e três parâmetros. Na área educacional é crescente o interesse pela aplicação de técnicas derivadas da TRI, já que essa metodologia vem sendo utilizada em processos qualitativos de avaliação psicológica e educacional. Porém, na grande parte das avaliações que empregam itens de múltipla escolha, é comum a redução das respostas em padrões de acerto ou erro para a utilização desses modelos Vendramini et al. (2000). Isso acontece devido à facilidade na interpretação de seus parâmetros. Contudo, interpretar os parâmetros dos itens para o MRN não é uma tarefa simples.

3.3 Função de Informação do Item

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

$$I_i(\theta) = \frac{[\frac{d}{d(\theta)}P_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}, \quad (7)$$

Em que,

- i) $I_i(\theta)$ é a “informação” fornecida pelo item i no nível de habilidade θ_i ;
- ii) $P_i(\theta) = P(X_{ij} = 1/\theta)$ e $Q_i(\theta) = 1 - P_i(\theta)$.

No caso do modelo logístico de três parâmetros, a equação pode ser escrita como:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i} \right]^2 \quad (8)$$

Essa equação mostra a importância que têm os três parâmetros sobre o montante de informação do item. Isto é, a informação é maior:

- i) quando b_i se aproxima de θ_i ;
- ii) quanto maior for o a_i ;
- iii) e quanto mais c_i se aproximar de 0.

3.4 Função de Informação do Teste (FIT)

A informação fornecida pelo teste é simplesmente a soma das informações fornecidas por cada item que compõe o mesmo:

$$I(\theta) = \sum_{i=1}^I I_i(\theta) \quad (9)$$

Outra maneira de representar essa função de informação do teste é através do erro-padrão de medida, chamado na TRI de erro-padrão de estimação, que é dado por:

$$EP(\theta) = \frac{1}{\sqrt{I(\theta)}}. \quad (10)$$

É importante notar que essas medidas de informação dependem do valor de θ . Assim, a amplitude do intervalo de confiança para θ dependerá também do seu valor.

Alguns exemplos de curvas características e de curvas de informação (traçado pontilhado) de itens com diferentes combinações de valores dos parâmetros a e b são apresentados na Figura 4.

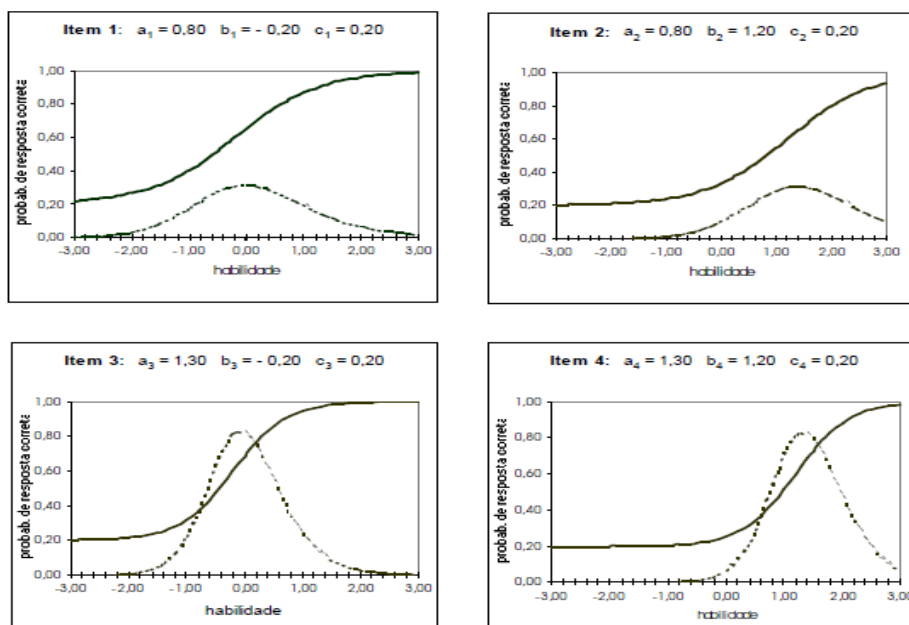
Comparando-se os itens 2 e 4 (e também os itens 1 e 3) pode-se perceber que os itens com maior valor do parâmetro a têm a curva característica com inclinação mais acentuada. A consequência disso é que a diferença entre as probabilidades de resposta correta de dois indivíduos com habilidades 2,00 e 1,00,

por exemplo, é maior no item 4 ($0,37=0,88-0,51$) do que no item 2 ($0,25=0,80-0,55$). Em outras palavras, o item 4 é mais apropriado para discriminar esses dois indivíduos do que o item 2. Por esse motivo é que o parâmetro a é denominado de parâmetro de discriminação (ou de inclinação) do item.

Por outro lado, comparando-se os itens um e dois (e também os itens três e quatro), pode-se perceber que os itens com maior valor do parâmetro b exigem uma habilidade maior para uma mesma probabilidade de resposta correta. Por exemplo, a habilidade requerida para uma probabilidade de resposta correta de 0,60 é igual a -0,20 no item um e igual a 1,20 no item dois. Isto é, o item dois é mais difícil do que o item um. Assim, o parâmetro b é denominado de parâmetro de dificuldade (ou de posição) do item.

Note que a cada item está associado um intervalo na escala de habilidade no qual o item tem maior poder de discriminação. Esse intervalo é definido em torno do valor do parâmetro b e está mostrado nos gráficos pelas curvas de informação (traçados pontilhados). Desse modo, a discriminação entre bons alunos é feita a partir de itens considerados difíceis e não de itens considerados fáceis.

Figura 4 – Curvas características e de informação de vários itens



Fonte: Andrade et al. (2000).

Apesar de receberem a mesma denominação da Teoria Clássica de Medida, o parâmetro de dificuldade do item não é medido por uma proporção (valor entre 0 e

1) e o parâmetro de discriminação não é uma correlação (valor entre -1 e 1). Na TRI, estes dois parâmetros podem, teoricamente, assumir qualquer valor real entre $-\infty$ e $+\infty$. Porém, como já foi dito, não se espera um valor negativo para o parâmetro a .

Na prática, as habilidades e os parâmetros dos itens são estimados a partir das respostas de um grupo de indivíduos submetidos a esses itens, mas uma vez estabelecida a escala de medida da habilidade, os valores dos parâmetros dos itens não mudam, isto é, seus valores são invariantes a diferentes grupos de respondentes, desde que os indivíduos desses grupos tenham suas habilidades medidas na mesma escala.

3.5 A Escala de Habilidade

Diferentemente da medida score em um teste com I questões do tipo certo/errado, que assume valores inteiros entre 0 e I , na TRI a habilidade pode teoricamente assumir qualquer valor real entre $-\infty$ e $+\infty$. Assim, precisa-se estabelecer uma origem e uma unidade de medida para a definição da escala.

Esses valores são escolhidos de modo a representar, respectivamente, o valor médio e o desvio-padrão das habilidades dos indivíduos da população em estudo. Para os gráficos mostrados anteriormente, utilizou-se a escala com média igual a 0 e desvio-padrão igual a 1, que será representada por escala (0,1). Essa escala é bastante utilizada pela TRI, e nesse caso, os valores do parâmetro b variam (tipicamente) entre -2 e +2. Com relação ao parâmetro a , espera-se valores entre 0 e +2, sendo que os valores mais apropriados de a seriam aqueles maiores do que 1.

Apesar da frequente utilização da escala (0,1), em termos práticos, não faz a menor diferença estabelecer-se esses valores ou outros quaisquer. O importante são as relações de ordem existentes entre seus pontos. Por exemplo, na escala (0,1) um indivíduo com habilidade 1,20 está 1,20 desvios-padrão acima da habilidade média. Esse mesmo indivíduo teria a habilidade 248, e consequentemente estaria também 1,20 desvios-padrão acima da habilidade média, se a escala utilizada para essa população fosse a escala (200, p. 40). Isso pode ser visto a partir da transformação de escala:

$$a(\theta - b) = \left(\frac{a}{40}\right)[(40 \times \theta + 200) - (40 \times b + 200)] = a^*(\theta^* - b^*), \quad (11)$$

em que $a(\infty - b)$ é a parte do modelo probabilístico proposto envolvida na transformação. Assim, tem-se que:

$$\text{i)} \quad \theta^* = 40x\theta + 200, \quad (12)$$

$$\text{ii)} \quad b^* = 40xb + 200, \quad (13)$$

$$\text{iii)} \quad a^* = a/40, \quad (14)$$

$$\text{iv)} \quad P(U_i = 1/\theta) = P(U_i = 1/\theta^*). \quad (15)$$

Por exemplo, os valores dos parâmetros a e b do item um, mostrado anteriormente, na escala (0,1) são, respectivamente, 0,80 e -0,20 e seus correspondentes na escala (200; 40) são, respectivamente, $0,02 = 0,80 / 40$ e $192 = 40 \times (-0,20) + 200$. Além disso, um indivíduo com habilidade $\theta = 1$; 00 medida na escala (0,1) tem sua habilidade representada por $\theta^* = 40 \times 1,00 + 200 = 240$ na escala (200; 40) e:

$$P(U_1 = 1/\theta = 1) = 0,20 + (1 - 0,20) \frac{1}{1 + e^{-1,7 \times 0,02 \times (1 - 0,20)}} \quad (16)$$

$$= 0,20 + (1 - 0,20) \frac{1}{1 + e^{-1,7 \times 0,02 \times (240 - 192)}}$$

$$= P(U_1 = 1/\theta^* = 240) = 0,87,$$

ou seja, a probabilidade de um indivíduo responder corretamente a um certo item é sempre a mesma, independentemente da escala utilizada para medir a sua habilidade, ou ainda, a habilidade de um indivíduo é invariante à escala de medida. Assim, não faz qualquer sentido querermos analisar itens a partir dos valores de seus parâmetros a e b sem conhecer a escala na qual eles foram determinados.

3.6 Suposições do Modelo: Unidimensionalidade e Independência Local

O modelo proposto pressupõe a unidimensionalidade do teste, isto é, a homogeneidade do conjunto de itens que supostamente devem estar medindo um único traço latente. Em outras palavras, deve haver apenas uma habilidade

responsável pela realização de todos os itens da prova. Parece claro que qualquer desempenho humano é sempre multideterminado ou multimotivado, dado que mais de um traço latente entra na execução de qualquer tarefa. Contudo, para satisfazer o postulado da unidimensionalidade, é suficiente admitir que haja uma habilidade dominante (um fator dominante) responsável pelo conjunto de itens. Esse fator é o que se supõe estar sendo medido pelo teste.

Tipicamente, a dimensionalidade do teste é verificada através da análise fatorial, feita a partir da matriz de correlações tetracóricas. Mislevy (1986) discute as deficiências da aplicação desse procedimento e sugere outro procedimento baseado no método de máxima verossimilhança.

Outra suposição do modelo é a chamada independência local ou independência condicional, a qual assume que para uma dada habilidade as respostas aos diferentes itens da prova são independentes. Esta suposição é fundamental para o processo de estimação dos parâmetros do modelo. Na realidade, como a unidimensionalidade implica independência local – ver Hambleton e Swaminathan (1991) –, tem-se somente uma e não duas suposições a serem verificadas. Assim, itens devem ser elaborados de modo a satisfazer a suposição de unidimensionalidade.

As vantagens da utilização da TRI dependem fundamentalmente da adequação (ajuste) dos modelos e seus pressupostos. Por exemplo, somente a partir de modelos com bom ajuste é que se pode garantir a obtenção de itens e habilidades invariantes.

3.7 Estimação: uma única população

A estimação dos parâmetros dos itens e das habilidades dos respondentes, é considerada uma das etapas mais importantes da TRI. A probabilidade de uma resposta correta a um determinado item depende somente da habilidade do indivíduo e dos parâmetros que caracterizam o item. Geralmente, ambos são desconhecidos. Apenas as respostas dos indivíduos aos itens do teste são conhecidas.

Entretanto, nos Modelos de Resposta ao Item temos um problema de estimação que envolve dois tipos de parâmetros, os parâmetro dos itens e as habilidades dos indivíduos. Então, do ponto de vista teórico, podemos dividir o

problema em três situações, quando já conhecemos os parâmetros dos itens, temos apenas que estimar as habilidades; se já conhecemos as habilidades dos respondentes, estaremos interessados apenas na estimação dos parâmetros dos itens e, por fim, a situação mais comum, em que desejamos estimar os parâmetros dos itens e as habilidades dos indivíduos simultaneamente. Na TRI, o processo de estimação dos parâmetros dos itens é conhecido como calibração.

Na situação em que desejamos estimar tanto os parâmetros dos itens, quanto as habilidades, há duas abordagens usuais: estimação conjunta, parâmetros dos itens e habilidades, ou em duas etapas, primeiro a estimação dos parâmetros dos itens e, posteriormente, das habilidades. No caso da estimação conjunta, é necessário estabelecer uma métrica para estimar os parâmetros. A necessidade de se estabelecer essa métrica consiste em um problema denominado falta de identificabilidade do modelo. A não identificabilidade ocorre também no MRN, pois diferentes valores dos parâmetros dos itens e dos traços latentes podem produzir a mesma probabilidade de alguma(s) categoria(s).

$$\begin{aligned}
 P\left(Y_{ijk} = 1 / \theta_j, \zeta_i\right) &= \frac{e^{a_{ik}(\theta_j - b_{ik})}}{\sum_{h=1}^{m_i} e^{a_{ih}(\theta_j - b_{ih})}} \quad (17) \\
 &= \frac{e^{a_{ik}(\theta_j - b_{ik}) \frac{\alpha}{\alpha}}}{\sum_{h=1}^{m_i} e^{a_{ih}(\theta_j - b_{ih}) \frac{\alpha}{\alpha}}} \\
 &= \frac{e^{\frac{a_{ik}}{\alpha}[\alpha\theta_j - \alpha b_{ik} + \beta - \beta]}}{\sum_{h=1}^{m_i} e^{\frac{a_{ih}}{\alpha}[\alpha\theta_j - \alpha b_{ih} + \beta - \beta]}} \\
 &= \frac{e^{\frac{a_{ik}}{\alpha}[(\alpha\theta_j + \beta) - (\alpha b_{ik} + \beta)]}}{\sum_{h=1}^{m_i} e^{\frac{a_{ih}}{\alpha}[(\alpha\theta_j + \beta) - (\alpha b_{ih} + \beta)]}} \\
 &= P\left(Y_{ijk} = 1 / \theta'_j, \zeta'_i\right).
 \end{aligned}$$

É interessante notar que no MRN, mesmo quando se fixam os traços latentes, tem-se o problema da falta de identificabilidade, ou seja, mesmo quando se

estabelece uma escala para os traços latentes, é necessário colocar uma restrição nos parâmetros dos itens.

3.7.1 Estimação dos Parâmetros dos Itens

Assim como nos modelos dicotômicos, o MRN possui algumas suposições que são importantes para o desenvolvimento do processo de estimação:

- i) S1 – as respostas oriundas de indivíduos diferentes são independentes;
- ii) S2 – os itens são respondidos de forma independente por cada indivíduo dado o seu traço latente;
- iii) S3 – a probabilidade de escolha da alternativa k do item i pelo indivíduo j pode ser descrita como.

$$P\left(Y_{ij} = y_{ij} / \theta_j, \zeta_i\right) = \prod_{k=1}^{m_i} P_{ijk}^{y_{ijk}} \quad (18)$$

em que P_{ijk} é uma função de resposta ao item relacionada com algum modelo politômico.

3.7.2 Estimação dos Traços Latentes

Levando-se em consideração que os parâmetros dos itens são conhecidos, ou estimados em uma etapa anterior, a estimação dos traços latentes é feita individualmente. A log-verossimilhança para um determinado indivíduo é dada por:

$$l(\theta_j) = \sum_{i=1}^I \sum_{k=1}^{m_i} y_{ijk} \ln P_{ijk}. \quad (19)$$

O EMV de θ_j é encontrado resolvendo o seguinte sistema de equações de estimação:

$$S(\theta_j) = \sum_{i=1}^I \sum_{k=1}^{m_i-1} \alpha_{ik} (P_{ij(k+1)} - Y_{ij(k+1)}). \quad (20)$$

Segundo Pereira (2004), se for de interesse medir tal traço latente faz-se necessário então criar uma escala de medida segundo a qual essa variável assumirá seus valores. Por inúmeras razões técnicas a definição da escala de medidas, o número na escala e a interpretabilidade da mesma em relação ao traço latente é muito difícil, necessitando de um formalismo maior. Para itens dicotômicos cada nível de habilidade existirá certa probabilidade que o respondente j com essa habilidade dará uma resposta correta ao item i .

3.8 Equalização

Andrade e Valle (1998) destacam uma característica que também pode ser considerada uma das grandes vantagens na utilização da TRI. Refere-se sobre a possibilidade de equalizar as habilidades dos indivíduos, sejam eles pertencentes à mesma população ou não, submetidos ao mesmo item ou conjunto de itens, cujos resultados permitem a comparação de seus desempenhos na atividade solicitada.

A equalização é um instrumento importante para se comparar o desempenho de sujeitos pertencentes ou não à mesma população. Para Andrade, Tavares e Valle (2000), a equalização garante a comparação entre itens e/ou níveis de habilidades, pois esse instrumento permite que os parâmetros de itens e/ou níveis de habilidades dos respondentes, de grupos diferentes, possam ser analisados e compreendidos em uma mesma métrica, ou seja, uma escala comum. Portanto,

Equalizar significa equiparar, tornar comparável, colocar os parâmetros dos itens provenientes de testes diferentes e traços latentes de respondentes de diferentes grupos na mesma métrica, tornando os itens e os respondentes comparáveis. Existem dois tipos de equalização: equalização via população, quando um único grupo de respondentes é submetido aos testes; via itens, quando grupos diferentes respondem testes diferentes com itens comuns entre eles. O segundo tipo de equalização pode ser realizado de dois modos: a posteriori e simultaneamente, através da utilização de modelos de grupos múltiplos. (ARAÚJO et al., 2009, p. 7).

Equalizar compreende a comparação de resultados, significa equiparar, tornar comparável, na TRI significa colocar parâmetros de itens vindos de provas distintas ou habilidades de respondentes de diferentes grupos, na mesma métrica, ou seja, numa escala comum, tornando os itens e/ou habilidades comparáveis (ANDRADE et al., 2000). Sem a equalização, não é possível afirmar, por exemplo, que a média de

desempenho de alunos em um Sistema de Avaliação sofreu mudanças de um período para outro.

É fundamental distinguir as diferenças de habilidade dos indivíduos avaliados, das diferenças de dificuldade dos instrumentos aplicados Rabello (2001). Essa técnica mostra eficiente quando se observa que a decisão baseada nos dados de qualquer sujeito deve ser a mesma independente da forma que ele respondeu e, ainda em situações em que não se quer tomar decisões a nível individual, as análises devem ser feitas levando-se em consideração os escores transformados pela equalização, que são comparáveis.

Há dois tipos de equalização: a equalização via população e a equalização via itens comuns. Ou seja, existem duas maneiras de colocar parâmetros, tanto de itens quanto de habilidades, numa mesma métrica (ANDRADE et al., 2000). Na primeira, se um único grupo de respondentes é submetido a provas diferentes, basta que todos os itens sejam calibrados conjuntamente para se ter a garantia de que todos estarão na mesma métrica. Na equalização via itens comuns, a garantia de que as populações envolvidas terão seus parâmetros em uma mesma escala será dada pelos itens comuns entre as populações que servirão de ligação entre elas.

3.8.1 Delineamento de grupos não equivalentes com itens comuns

Para realizar uma equalização é preciso planejar um delineamento. Como exemplo desse delineamento, pode-se valer de uma aplicação contendo duas formas, X e Y, contendo um conjunto de itens comuns. Para Hambleton e Swaminathan (1995), esse delineamento é chamado de delineamento do teste âncora. Para esse método, existem duas variações possíveis:

- i) os itens comuns contribuem para o cálculo do escore do indivíduo no teste, sendo chamado de itens internos;
- ii) os itens comuns não contribuem para o cálculo do escore do indivíduo, sendo denominado de itens externos.

Os grupos de itens internos e externos são geralmente aplicados com tempo marcado separadamente.

Para Yang e Houang (1996) independentemente do método de equalização usado em um delineamento que usa itens comuns, o método será mais preciso à medida que aumentar o número de itens comuns. Segundo Andrade (1999) e Yang e Houang (1996), uma referência para o número mínimo de itens comuns é 20% do tamanho total de cada uma das provas a ser equalizada. Além disso, Yang e Houang afirmam que independentemente do delineamento utilizado, os testes a serem equalizados devem possuir 35 itens no mínimo.

3.8.2 Diferentes tipos de equalização

Existem dois tipos de equalização: equalização via população, quando um único grupo de respondentes é submetido aos testes; via itens, quando grupos diferentes respondem testes diferentes com itens comuns entre eles. O segundo tipo de equalização pode ser realizado de dois modos: a posteriori e simultaneamente, através da utilização de modelos de grupos múltiplos (ANDRADE et al., 2000).

3.8.2.1 Um único grupo fazendo uma única prova

Esse é o caso trivial, em que se aplicam diretamente os modelos matemáticos e os métodos de estimação descritos anteriormente. Ou seja, não é necessário nenhum tipo de equalização.

3.8.2.2 Um único grupo fazendo duas provas totalmente distintas

Tem-se equalização via população. Para resolvê-la, basta que todos os itens de ambas as provas sejam calibrados simultaneamente. O fato de todos os indivíduos representarem uma amostra aleatória de uma mesma população é que garante que todos os parâmetros envolvidos estarão numa mesma escala.

3.8.2.3 Um único grupo fazendo duas provas parcialmente distintas

Esse caso é bastante semelhante ao caso anterior, podendo fazer também a equalização via população.

3.8.2.4 Dois grupos fazendo uma única prova

É exemplo de equalização via itens comuns. Como as duas populações fazem exatamente a mesma prova, basta que os itens sejam calibrados, utilizando-se as respostas dos respondentes de ambos os grupos, simultaneamente.

3.8.2.5 Dois grupos fazendo duas provas totalmente distintas

Esse caso não pode ser resolvido pela TRI. É possível calibrar separadamente os itens das duas provas, porém não é possível fazer nenhum tipo de comparação entre os resultados obtidos, uma vez que eles estarão em métricas diferentes. Então não faz sentido comparar os resultados desses dois grupos.

3.8.2.6 Dois grupos fazendo duas provas parcialmente distintas

Nessa equalização via itens comuns, pode ser considerado o maior avanço da TRI sobre a Teoria Clássica. O uso de itens comuns entre provas distintas aplicadas a populações distintas permite que todos os parâmetros estejam na mesma escala ao final dos processos de estimação, possibilitando comparações e a construção de “escalas de conhecimento” interpretáveis, que são de grande importância na área educacional.

3.9 Diferentes problemas de estimação

A estimação dos parâmetros dos itens e dos desempenhos é uma das etapas mais importantes da TRI. Segundo Andrade, Tavares e Valle (2000, p. 27), geralmente essa estimação de parâmetros, ou calibração, é feita pelo Método da Máxima Verossimilhança, através de algum processo iterativo. Assim como, as equações matemáticas, utilizadas para os métodos de estimação, estão demonstradas nas obras de Andrade et al. (2000), nas quais pode-se verificar que a probabilidade de uma resposta correta a um determinado item depende do desempenho do respondente e dos parâmetros que caracterizam o item, em geral ambos desconhecidos, sendo conhecidas somente as respostas dos respondentes, ao conjunto de itens.

3.9.1 Quando todos os itens são novos

Nesse caso, deseja-se calibrar o conjunto completo de itens. Esse é o caso trivial e para resolver esse problema, basta utilizar uma das técnicas de estimação descritas anteriormente.

3.9.2 Quando todos os itens já estão calibrados

Esse é o caso em que todos os itens já foram calibrados anteriormente, ou seja, deseja-se apenas estimar as habilidades dos respondentes. Esse também é um caso bastante frequente na TRI, devido ao impulso que essa teoria deu na criação de bancos de itens. Esses bancos são formados por conjuntos de itens que já foram testados e calibrados a partir de um número significativo de indivíduos de uma dada população. Dessa forma, assume-se que os parâmetros desses itens já são “conhecidos”, ou seja, assume-se que são conhecidos os verdadeiros valores dos parâmetros desses itens e assim, sempre que se desejar, podem-se aplicar novamente alguns desses itens do banco a outros indivíduos e assim estimar apenas suas habilidades que estarão sempre na mesma métrica dos parâmetros dos itens.

Quando se “constrói” um banco de itens, uma informação fundamental é a escala em que aqueles itens foram calibrados, visto que as habilidades de indivíduos que serão estimadas futuramente a partir daqueles itens estarão nessa mesma métrica e, portanto, quaisquer comparações diretas só poderão ser feitas com outros indivíduos que também tenham suas habilidades nessa escala.

Dessa forma, para resolver esse problema, basta utilizar um dos processos de estimação das habilidades dos indivíduos quando os parâmetros dos itens já são conhecidos.

3.9.3 Quando alguns itens são novos e outros já estão calibrados

Nesse caso, deseja-se calibrar alguns itens e manter os parâmetros de outros, que já foram calibrados anteriormente. Esse caso também está ligado à criação de bancos de itens, uma vez que um banco de itens está continuamente em

formação, ou seja, pode-se estar interessado em acrescentar novos itens ou retirar itens ao conjunto que já se encontra no banco.

Para resolver esse problema, deve-se definir uma das populações como sendo a referência, e então, as demais populações serão posicionadas em relação a ela, a fim de evitar problemas de indeterminação de escala.

3.10 A Escala de Habilidade

A escala é uma ferramenta utilizada para sistematizar informações. Nela, utiliza-se o conceito de medida que atribui números aos fenômenos naturais. As habilidades dos estudantes são fenômenos naturais observáveis para os quais podem ser atribuídas quantidades. Uma escala é capaz de representar em que e quanto mais grupos de estudantes construíram competências e habilidades que outros grupos de estudantes ainda não construíram.

Quando se fala em métrica, significa se referir ao tipo de escala utilizada para medir um dado fenômeno, ou seja, se um indivíduo que obtém nota nove é considerado excelente numa prova de desempenho, então supõe-se que a métrica utilizada seja uma escala que vai de 0 a 10, pois se a escala fosse de 0 a 100, então a nota representaria péssimo desempenho. Assim, é de fundamental importância saber a métrica utilizada para poder entender o significado do valor atribuído.

A escala da habilidade é uma escala arbitrária em que o importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude. O parâmetro c não depende da escala, pois se trata de uma probabilidade, e como tal, assume sempre valores entre 0 e 1. O parâmetro b representa a habilidade necessária para uma probabilidade de acerto casual igual a $(1 + c)/2$. Assim, quanto maior o valor de b , mais difícil é o item, e vice-versa.

Na TRI, a habilidade pode assumir teoricamente qualquer valor real entre $-\infty$ e $+\infty$. Daí, é necessário estabelecer uma origem e uma unidade de medida para a definição da escala. Esses valores são escolhidos de modo a representar, respectivamente, o valor médio e o desvio-padrão das habilidades dos indivíduos da população em estudo. Em muitos casos, utiliza-se a escala com média igual a 0 e desvio-padrão igual a 1, que é representada por escala (0,1). Essa escala é bastante utilizada pela TRI e, nesse caso, os valores do parâmetro b variam entre -2 e $+2$.

Com relação ao parâmetro a , espera-se valores entre 0 e +2, sendo que os valores mais apropriados de a seriam aqueles maiores que 1 (ANDRADE et al., 2000).

3.11 Trabalhos Relacionados

Nas duas últimas décadas, considerando as constantes inovações produzidas pelos avanços tecnológicos, é indiscutível repensar os Sistemas Avaliativos. Partindo do princípio de que a avaliação contribui para auxiliar o ensino, orientar a aprendizagem, obter informações sobre os alunos, os professores e a instituição, bem como certificar os resultados ao final de um período letivo. Dessa forma, buscam-se estratégias condizentes com o momento atual em que “a avaliação educacional vem sofrendo uma transformação radical com a mudança da cultura da prova para a cultura da avaliação” (VIANNA, 1997, p. 36). Compreende-se que a grande mudança que se coloca hoje para a avaliação é se estabelecer como parte do processo de ensino e de aprendizagem, de forma integrada.

Para tanto, têm surgido um bom número de ambientes e ferramentas específicas que tratam de forma única o aspecto de avaliação dos alunos. Para melhor compreensão, destacam-se três pesquisas que estão fortemente ligadas ao tema desta proposta, objetivando uma melhor visualização para a diversidade de enfoques que a pesquisa em avaliação educacional apresenta.

3.11.1 AVALWEB – Sistema Interativo para Gerência de Questões e Aplicação de Avaliações na Web

Desenvolvido pela Universidade Federal do Rio Grande do Sul, possui como foco o gerenciamento de questões, que inclui: o atendimento de requisições dos professores para elaboração de avaliações; escolha do nível de dificuldade das questões que comporão a avaliação; e uso de estratégias interativas para aplicação de provas.

De acordo com Cardoso (2001), em se tratando do ciclo de vida, o AVALWEB possui enfoque principal na fase de preparação, permitindo ao professor, a partir de certos critérios quanto ao nível de dificuldade e tipos de questões, gerar avaliações personalizadas. Além do nível de dificuldade da questão informada pelo professor, o

sistema calcula um nível próprio baseado em número de acertos/erros, que será posteriormente sugerido ao professor.

Caso o professor opte por uma avaliação do tipo adaptativa (a geração de questões dependerá das respostas dos alunos, em questões anteriores), deverá determinar se, quanto ao nível de dificuldade, essa será progressivamente maior ou menor.

Ressalta-se que o AVALWEB possui um módulo de autoavaliação possibilitando retorno imediato ao aluno, não o classificando com relação aos outros estudantes.

3.11.2 JAVAL – Ambiente para Avaliação Remota para Ensino à Distância Empregando Agentes Embarcados

Ambiente desenvolvido em Java, cuja proposta foi selecionada pela CAPES no Programa de Apoio a Pesquisa em Educação a Distância (PAPED-2001). Permite a criação de avaliações e suas questões, de tipo: certo-errado, múltipla-escolha, associação e lacuna; possuindo como principal objetivo o monitoramento do aluno durante a realização dessas avaliações, detectando situações de dúvida e iniciando procedimentos de auxílio.

Quanto ao ciclo de vida, segundo Emiliano (2002), possui seu foco na fase de avaliação e mais especificamente no aspecto rastreamento do aluno. Os agentes embarcados acompanham o aluno durante a solução de sua avaliação, monitorando: aspectos comportamentais, como minimizações de janela, modificação de respostas, etc; e aspectos temporais, em que monitoram o tempo de realização das questões. Quando algum desses eventos ocorre fora da margem prevista pelo professor, o agente disponibiliza ao aluno recursos de ajuda como páginas da Web, arquivos de áudio, vídeo, lista de discussão, etc.; todos previamente definidos pelo professor.

Finalmente, após a realização da avaliação, o sistema disponibiliza um relatório ao professor, contendo informações do tipo: questões resolvidas fora do tempo previsto, questões não resolvidas, nota do aluno, etc.

3.11.3 Modelagem para o Software Virtual-Taneb Baseado na Teoria da Resposta ao Item para Avaliar o Rendimento dos Alunos

Para Veras (2010), esse *software* proporcionará um suporte aos professores e pedagogos na avaliação do conhecimento de Matemática dos alunos da quarta série do ensino fundamental, especificamente no domínio da geometria. Através desse, será determinado o nível de desempenho dos alunos, a partir das questões respondidas corretamente, aplicando a Teoria da Resposta ao Item para avaliar as respostas.

A Interface com o usuário auxiliará o estudante na realização de sua prova e, posteriormente, permitindo que professores e pedagogos possam avaliar o nível de conhecimento desses alunos sobre uma determinada área. Segundo o autor, essa interação possibilita que através dessas respostas, possa-se fazer uma análise e, através do resultado da análise, inferir a qualidade do ensino fundamental na área de Matemática em uma determinada escola. Dessa forma, a GUI é responsável por permitir o contato do aluno com as provas formuladas a partir da seleção das questões armazenadas em um banco de dados através de agente de software. E, também, possibilita o contato dos professores e pedagogos com os rendimentos dos alunos através dos resultados obtidos conforme os modelos matemáticos baseados na Teoria da Resposta ao Item.

É importante salientar que a aplicação da Teoria da Resposta ao Item nesse processo é fundamental, pois através dela é possível determinar o modelo matemático a ser adotado na avaliação da capacidade de conhecimento do aluno em relação à resolução de provas sobre a área de Matemática. Portanto, o Virtual-TANEB levará em consideração o modelo logístico de três parâmetros, a fim de possibilitar a construção de uma estimativa de uma única população analisando a afinidade e as habilidades contidas em cada aluno.

As principais características distintas do Projeto Virtual-TANEB em relação à maneira como é feita a prova atualmente pelo SAEB são:

- i) Estabelecer mecanismos computacionais, utilizando Inteligência Artificial para melhorar e aperfeiçoar a avaliação do nível de conhecimento dos alunos em relação à matemática da quarta série do ensino fundamental;
- ii) Automatizar o processo de avaliação da educação básica;
- iii) Reduzir o tempo de resposta atual para a execução, análise e resultado do diagnóstico da prova do SAEB;

- iv) Permitir de forma paralela e distribuída a aplicação da prova a várias escolas, através de um portal na Internet.

Visualização panorâmica das perspectivas pedagógicas dos Sistemas AVALWEB, JAVAL e VIRTUAL-TANEB, destacando as principais concepções de ensino e de aprendizagem, bem como linguagem utilizada.

Quadro 1 – Visualização panorâmica das perspectivas pedagógicas dos Sistemas AVALWEB, JAVAL e VIRTUAL-TANEB

PRINCIPAIS CARACTERISICAS	AVALWEB	JAVAL	VIRTUAL-TANEB
Teoria	Gerenciamento de questões e aplicação de avaliações via Web com base em requisições de professores.	Os ambientes de Ensino a Distância (EAD) dinâmico e assíncrono.	Teoria da Resposta ao Item (TRI).
Principais teóricos	Bloom	Bloom	Castro, Andrade, Tavares e Valle.
Implicações para a aprendizagem	Possui um módulo de autoavaliação, com retorno imediato, dando ênfase mais no processo de ensino/aprendizagem do que na avaliação.	Medidas que permitam ao professor acompanhar o aluno durante as avaliações.	Interação através das respostas, podendo-se inferir na qualidade do ensino.
Implicações para o ensino	Fornecer subsídios para ajudar o processo de avaliação através de dados estatístico.	O uso de agentes de <i>software</i> embarcados no ambiente do aluno, minimiza o efeito de dispersão espacial e temporal destes ambientes,	Proporcionará um suporte aos professores e pedagogos na avaliação do conhecimento de Matemática dos alunos da quarta série do ensino

		obtendo <i>feedback</i> ao professor quanto à sua elaboração.	fundamental.
Implicações para a avaliação	A avaliação inclui uma grande variedade de dados, superior a o rotineiro exame escrito final.	O aluno pode ser avaliado de forma somativa, pela obtenção de estatísticas de acertos e erros.	Professores e os pedagogos realizarão um diagnóstico após as atividades de resolução dos itens.
Questão de validação	Com base na estrutura dos documentos e nos contextos, é definida a estrutura de navegação do sistema, especificando o relacionamento entre documentos e os caminhos de acesso às funções.	Um estudo de caso no Centro Universitário do Pará, durante o 2º semestre do ano 2000.	Determinará o nível de desempenho dos alunos, aplicando a Teoria da Resposta ao Item para avaliar as respostas.

Fonte: A autora

Constata-se a necessidade de repensar os processos de avaliação, os avanços tecnológicos têm possibilitado a otimização dos instrumentos de avaliação, permitindo sua implementação de maneira dinâmica, sistêmica e integrada.

4 MATERIAL E MÉTODO

Para concretização dessa pesquisa, faz-se necessário o uso de uma metodologia que, na visão de Richardson (2010), são procedimentos e regras utilizadas por determinado método. De acordo com Pacheco Júnior et al. (2007), através do método pode-se chegar a determinado objetivo. Desse modo, compreende-se que método é a forma de proceder ao longo de um caminho, seguindo os diferentes processos necessários para atingir um determinado fim. Nesta seção, delinea-se o aporte metodológico utilizado para implementação desta pesquisa, bem como o desenvolvimento da sistemática proposta e análise estatística da situação.

4.1 Caracterização da Pesquisa

Essa pesquisa é do tipo teórico-aplicada, pois visa propor, aplicar e validar uma sistemática para a implantação de um sistema de Avaliação utilizando a TRI. Gil (2010) considera que o objetivo da pesquisa aplicada é gerar conhecimentos com o propósito de aplicá-los em uma determinada situação.

Entretanto, é uma pesquisa é do tipo quantitativa, pois utiliza perguntas baseadas em instrumento, dados de desempenho e análise estatística para o tratamento dos dados (CRESWELL, 2007). Richardson (2010) considera que o método quantitativo é amplamente utilizado na condução das pesquisas, representando a intenção de garantir a precisão dos resultados e evitar distorções de análise e interpretação, possibilitando uma margem de segurança quanto às inferências.

Assim sendo, na visão de Gil (2010), temos uma pesquisa exploratória porque utiliza pesquisa bibliográfica. Gil (1999, p. 43) ressalta que “pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar visão geral, de tipo aproximativo, acerca de determinado fato”. Portanto, visa proporcionar maior familiaridade com a questão, com vistas a torná-la mais explícita, facilitando a construção de hipóteses. Com o objetivo de proporcionar melhor entendimento do tema estudado. Realizou-se uma pesquisa bibliográfica sobre TRI, na qual a mesma é utilizada na elaboração de uma sistemática para a implementação de uma ferramenta avaliativa, levando em consideração a problemática apresentada.

Dessa forma, essa pesquisa compreende os procedimentos técnicos de pesquisa bibliográfica, pelo exaustivo levantamento de referências relacionadas à TRI, utilizadas para a elaboração da sistemática proposta. Também são utilizados procedimentos experimentais, pela utilização de simulações e análises de dados. Além de caracterizar a situação dos alunos do curso Licenciatura em Matemática do CESC/UEMA, em que foi aplicado o teste.

4.2 Etapas da Pesquisa

Essa pesquisa foi basicamente constituída por uma parte teórica e por uma parte prática, caracterizada pela utilização da sistemática proposta, aplicada em uma situação específica envolvendo alunos do curso Licenciatura em Matemática do CESC/UEMA.

Na parte teórica, inicialmente, foi feito um breve levantamento bibliográfico sobre Avaliação em Larga Escala e a TRI. Em seguida, foi realizado levantamento bibliográfico sobre o *Software* jMetrik, incluindo testes, relatórios técnicos e artigos de periódico. Na sequência, foram apresentados alguns trabalhos relacionados a essa pesquisa.

Na parte prática, foi desenvolvido um Sistema de Avaliação composto por uma ferramenta de autoria, ferramenta de aplicação e corretor de provas com os seus respectivos casos de uso.

Foi realizado um estudo, envolvendo os alunos do curso Licenciatura em Matemática, que compreende uma população de 150 alunos. A amostra se constitui por 45 alunos do 6º período do curso supracitado, realizando uma prova de 25 itens, retirados de um banco de itens do ENADE 2005/2008/2011, e aplicadas em junho de 2013.

Nesse estudo, foi feita a análise dos itens das provas de Matemática aplicada aos alunos do 6º período do curso Licenciatura em Matemática do CESC/UEMA 2013, bem como, das habilidades dos indivíduos referentes à disciplina. Ressalta-se que 45 alunos participaram dessa avaliação. As provas de Matemática foram constituídas de 25 itens de múltipla escolha. De acordo com temas a seguir:

Quadro 2 – Itens Analisados

Item	Descrição do Item
01	Operações envolvendo números racionais e área de figuras planas
02	Resolução de equação quadrática e elementos de Geometria
03	Números decimais e sua forma racional
04	Modelagem e solução de problemas do 1º grau
05	Solução de problemas do 1º grau envolvendo figuras espaciais
06	Problemas envolvendo porcentagem
07	Problemas envolvendo operações fundamentais
08	Investigação de dados e tabelas envolvendo números primos
09	Funções exponenciais e suas variações
10	Problemas envolvendo equações do 2º grau
11	Construções geométricas
12	Igualdade de números irracionais
13	Mudança de linguagem escrita para a linguagem Matemática
14	Soluções e discussões de Sistemas Lineares
15	Composições de funções
16	Limites de Função
17	Razão entre áreas de figuras planas
18	Números complexos na forma polar
19	Polinômios e porcentagem
20	Aplicação da Integral definida
21	Funções Diferenciáveis em $\mathbb{R}^3$
22	Estudo das Cônicas
23	Teoria dos Números
24	Aplicação da Derivada em funções reais
25	Derivadas Parciais

Fonte: A autora.

A análise dos dados foi feita baseada na Teoria da Resposta ao Item (TRI). Para a estimação dos parâmetros do modelo da TRI, foi utilizado um programa computacional visto que existe uma grande quantidade de dados que exigem compilação e também pela complexidade das operações. Na análise, poderiam ser utilizados os *softwares* TESTFACT (BOCK et al., 2003) para a análise da dimensionalidade, BILOG-MG (TOIT, 2003) para a calibração dos itens entre outros, porém, optou-se pelo uso do jMetrik, pois é específico para TRI. Além de ser *freeware* (de uso livre), *open source* (código fonte aberto) e feito em JAVA.

Figura 5 – Captura de tela com listagem das respostas dos alunos às respostas do teste aplicado

id	nome	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15	q16	q17	q18	q19	q20	q21	q22	q23	q24	q25
1,0000	aluno1	C	D	E	C	D	C	D	E	E	D	E	C	A	E	C	E	B	D	A	A	B	B	E	E	E
2,0000	aluno2	C	B	A	B	D	C	A	B	D	E	E	D	E	E	B	D	C	C	C	C	D	A	D	C	D
3,0000	aluno3	D	E	D	E	D	E	A	C	A	D	E	E	D	A	E	E	E	E	D	A	E	B	D	C	E
4,0000	aluno4	A	B	C	A	B	B	A	A	E	C	D	A	D	E	C	C	B	B	D	A	C	D	A	B	C
5,0000	aluno5	C	C	D	B	B	C	A	E	C	C	B	A	B	C	C	D	A	B	C	B	C	C	D	B	B
6,0000	aluno6	D	D	E	C	E	C	E	B	B	D	E	B	C	E	C	E	D	C	B	C	B	E	C	E	A
7,0000	aluno7	E	D	B	E	E	E	C	A	C	B	C	E	A	B	B	E	B	D	C	C	D	E	D	A	C
8,0000	aluno8	A	D	E	A	B	D	A	E	D	C	D	E	A	C	E	E	A	B	B	A	D	E	E	A	C
9,0000	aluno9	C	C	B	C	D	E	A	C	E	A	A	D	E	D	A	A	D	E	D	D	D	C	C	E	E
10,0000	aluno10	C	C	A	B	D	C	A	A	E	A	E	A	E	D	D	E	D	C	B	D	A	A	D	B	C
11,0000	aluno11	A	B	E	D	B	E	A	A	D	C	E	E	B	C	A	C	E	D	B	A	C	A	E	B	B
12,0000	aluno12	A	D	A	E	E	E	A	C	E	D	D	B	C	D	E	B	B	A	B	A	D	E	D	A	D
13,0000	aluno13	B	D	E	E	E	E	A	E	D	C	E	C	D	E	C	C	B	C	B	B	C	B	E	C	C
14,0000	aluno14	D	E	C	B	B	E	C	C	B	E	D	E	B	A	A	E	D	B	D	A	C	C	C	A	C
15,0000	aluno15	D	C	B	E	D	E	A	D	A	C	A	C	D	E	D	E	A	C	B	D	B	C	B	E	B
16,0000	aluno16	A	A	C	B	D	A	A	C	D	C	C	E	C	D	C	B	E	D	A	B	C	B	D	E	D
17,0000	aluno17	C	D	E	B	D	E	E	A	A	D	D	D	E	B	E	D	A	D	B	D	D	B	D	E	D
18,0000	aluno18	D	D	E	B	D	E	C	A	C	C	E	D	E	E	B	D	A	E	A	D	B	A	B	D	C
19,0000	aluno19	C	C	D	B	D	C	C	A	D	A	D	C	A	C	D	E	E	D	C	D	C	C	E	C	B
20,0000	aluno20	D	A	B	B	D	B	A	C	D	D	D	B	E	D	A	C	A	D	B	B	B	D	B	A	A
21,0000	aluno21	E	D	C	A	B	C	D	D	D	E	A	D	D	D	B	E	B	B	C	A	C	A	E	D	B
22,0000	aluno22	E	B	E	E	B	E	A	A	B	C	E	E	B	B	A	E	A	B	D	A	B	B	C	C	A
23,0000	aluno23	D	C	C	E	A	E	A	C	D	C	C	B	D	C	E	E	A	E	B	B	B	C	A	A	A
24,0000	aluno24	C	D	D	A	D	C	A	C	E	C	E	A	B	C	E	A	A	A	A	A	D	C	B	E	B
25,0000	aluno25	D	B	B	B	D	E	A	E	E	D	E	A	E	D	D	B	C	C	C	A	A	C	D	B	D
26,0000	aluno26	B	D	E	B	B	C	A	D	E	C	E	C	C	E	B	E	A	D	B	D	D	E	D	A	C
27,0000	aluno27	A	D	D	B	D	C	A	A	A	D	E	D	C	A	C	B	D	B	A	A	E	E	C	B	C
28,0000	aluno28	A	E	D	E	D	E	A	C	B	A	E	E	C	A	A	E	A	A	C	A	B	A	A	C	C
29,0000	aluno29	D	C	D	C	B	E	C	A	E	A	E	E	C	C	E	D	C	C	A	D	D	D	C	D	D
30,0000	aluno30	C	A	E	C	D	E	A	C	E	C	E	E	A	A	A	E	A	C	B	D	A	E	D	A	C
31,0000	aluno31	A	C	E	D	D	E	A	C	C	C	C	B	D	E	E	E	B	D	C	A	C	D	A	C	D
32,0000	aluno32	A	D	C	E	B	C	D	E	A	A	E	B	D	A	A	D	B	A	D	B	B	B	E	A	A
33,0000	aluno33	C	D	E	B	D	E	A	A	C	C	C	A	D	B	B	E	D	C	B	A	E	C	C	D	E
34,0000	aluno34	C	D	C	D	D	E	D	C	B	C	E	D	C	D	B	B	B	C	A	D	B	A	C	E	C
35,0000	aluno35	C	A	D	B	E	E	A	B	B	D	E	C	A	E	C	D	B	C	B	D	C	B	D	B	C
36,0000	aluno36	C	A	C	E	D	D	A	A	D	C	E	E	B	E	A	E	A	C	A	D	B	E	A	E	C
37,0000	aluno37	B	C	D	E	A	E	B	C	D	E	E	B	E	D	E	C	D	C	C	B	D	D	D	C	D
38,0000	aluno38	D	C	A	B	D	D	C	C	A	C	A	E	C	B	E	C	B	A	D	E	C	A	C	A	D
39,0000	aluno39	A	D	E	B	D	D	A	B	D	C	E	C	A	B	A	C	D	E	C	D	E	E	C	B	E
40,0000	aluno40	B	D	E	D	D	E	A	A	D	C	C	E	C	E	E	A	B	E	A	B	B	E	D	B	B
41,0000	aluno41	A	A	C	B	B	E	C	A	E	C	E	D	D	D	D	A	E	D	D	E	B	A	B	E	C
42,0000	aluno42	A	E	A	B	D	E	A	A	A	D	E	D	D	A	C	E	C	B	C	A	B	D	B	D	C
43,0000	aluno43	B	D	E	A	B	C	A	A	B	C	B	A	A	D	B	E	B	C	A	A	D	A	C	B	D
44,0000	aluno44	C	C	D	E	B	E	B	C	E	C	E	E	E	D	A	E	A	B	D	A	B	B	B	E	A
45,0000	aluno45	C	C	B	A	D	E	B	D	C	C	D	E	D	D	C	B	B	D	D	C	C	C	E	E	A

Fonte: Software jMetrik.

Nesta pesquisa, utilizou-se a ferramenta jMetrik para analisar os dados do experimento planejado para a “turma teste 1”, acima encontra-se o recorte da tela da ferramenta com o banco de dados das respostas dos alunos.

O jMetrik é uma aplicação de software livre e de código fonte aberto para análise psicométrica. Possui uma interface amigável, banco de dados integrado, e uma variedade de procedimentos estatísticos e gráficos. O jMetrik é um instrumento que combina uma variedade de procedimentos psicométricos em um aplicativo com uma interface amigável para as pessoas que o utilizam. Através dele é possível lidar com uma grande quantidade de dados, além de fornecer uma variedade de técnicas estatísticas para análise.

Foi criado por J. Patrik Meyer, professor e pesquisador da Universidade de Virginia, que recebeu o prêmio Bradley Hanson do Conselho Nacional de Medição em Educação (NCME), em 2010, por suas contribuições no campo da avaliação educacional. A versão mais recente é a 3.0, trata-se de uma revisão significativa do jMetrik que não só melhora a experiência do usuário, mas também adiciona novas funcionalidades para análise e gerenciamento de dados. Algumas das mudanças mais notáveis incluem a capacidade de:

- i) Criar novas tabelas, selecionando subconjuntos de variáveis;
- ii) Criar novas tabelas, selecionando subconjuntos dos examinandos;
- iii) Salvar saída de análise de itens para uma nova tabela no banco de dados;
- iv) Salvar saída da análise DIF para uma nova tabela no banco de dados;
- v) Salvar gráficos no disco como png ou jpg;
- vi) Automaticamente salvar várias parcelas de curvas características para disco como jpg;
- vii) Gerar automaticamente comandos de itens de pontuação a partir da interface;
- viii) Criar a sua própria pontuação correspondente para a análise de DIF;
- ix) Criar a sua própria estimativa de traço latente para as curvas características não paramétricas;
- x) Produzir curvas não paramétricas mais rapidamente, com menos esgotamento nos recursos do sistema.

São inúmeras as opções de escala de testes, apresentando facilidades de uso. O usuário pode rapidamente converter dados e especificar restrições no mínimo, máxima, e os pontos de precisão, bem como a conversão de uma

transformação linear personalizado. O programa permite renomear as novas variáveis marcadas, no entanto, apenas uma transformação pode ocorrer em uma execução e uma vez transformado, não há capacidade de renomear variáveis.

A escala de teste e recursos no jMetrik são úteis, fáceis de usar e entender. Portanto, as análises dos TCT fornecido pelo programa apresenta uma interface simples de *point-and-click* semelhante ao SPSS e Excel, mais fornece a tão necessária análise psicométrica em um programa que, até então, não está disponível para o especialista em psicometria.

De acordo com Gotzmann (2012), o programa jMetrik permite um sistema mais integrado para realizar análises psicométricas para fins de pesquisa e operações sem o custo associado com alguma outro programa de software psicométrico.

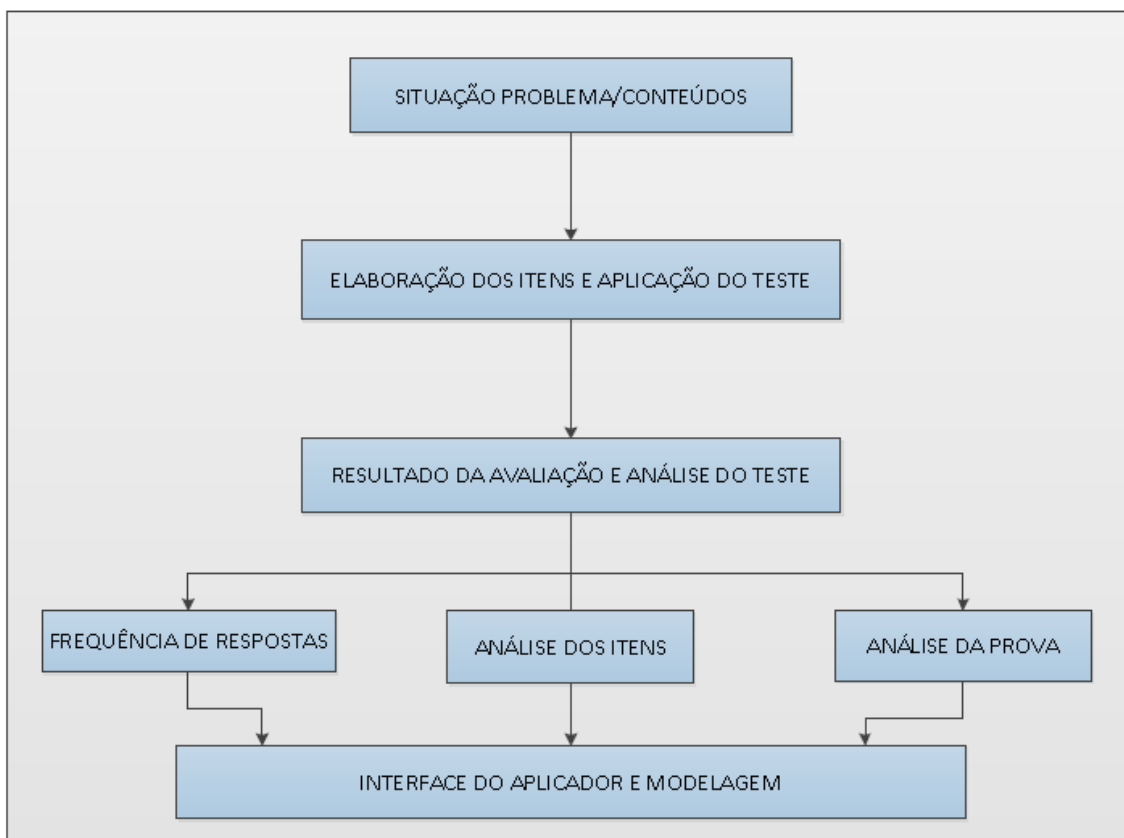
As análises da TRI, fornecidas pelo jMetrik, de acordo com Meyer (2010), incluem calibração de Rasch Crédito parcial e Escala Modelo de Avaliação. Esse programa também tem opções para a calibração, itens e pessoas, que incluem a convergência, marcando falta de dados, a fixação de valores de parâmetros dos itens, economizando estimativas teta e estatísticas de ajuste pessoal. Opções básicas de convergência estão disponíveis, tais como o número de iterações, e os critérios de convergência, mas faltam opções no *software*, tais como método de estimação.

As estatísticas da TRI são salvas no arquivo de dados, juntamente com as fontes de dados originais, que podem ser exportados para um arquivo de texto, para melhor compreensão do usuário.

Com ênfase nessas melhorias, evidencia-se a presença de novos recursos para a TRI, aumentando a capacidade do programa em estimar coeficientes de ligação com os testes de formato misto, usando qualquer combinação de Rasch, tais como: logística de dois parâmetros; logística de três parâmetros; o crédito parcial generalizado de crédito parcial ou modelos de resposta gradual.

Para o uso mais efetivo e eficaz da ferramenta, é fundamental que se intensifiquem a divulgação e a sistematização de conhecimentos sobre o jMetrik de forma que não se limitem a uma compreensão restrita, de modo que os usuários de todos os níveis de conhecimento possam tirar proveito em inúmeras análises psicométricas, principalmente na Avaliação Educacional.

Figura 6 – Detalhamento dos dados do teste



Fonte: A autora

A sistemática para detalhamento da pesquisa consiste em quatro etapas, conforme a Figura 6, que englobam todo o processo para a implantação e modelagem do sistema desde a sua definição até a sua efetiva aplicação. Segue a descrição de cada etapa:

A Etapa 1 – Situação problema / conteúdos, apresenta a situação a ser analisada, considerando a grande dificuldade que o Curso Licenciatura em Matemática vem se deparando nos últimos anos junto ao MEC, através de relatórios realizados pelo INEP.

A ideia da aplicação de uma avaliação do conhecimento do conteúdo específico do curso Licenciatura em Matemática, ministrado no CESC/UEMA nasce a partir do reconhecimento da realidade imposta pelo MEC, obrigando os estudantes do ensino superior a submeter-se à avaliação do ENADE a cada ciclo de três anos. Seu resultado poderá servir de base para a determinação de ações de ajuste a serem adotadas para a correção de qualquer desvio ou deficiência identificada no decorrer do curso superior, e antes da ocorrência da avaliação do ENADE.

Inicialmente propõe-se a realização da avaliação de Matemática aos alunos matriculados no 6º período do Curso de Matemática do CESC/UEMA. Os itens da prova foram retirados de um banco de itens, formado pelo conjunto de itens das provas do ENADE 2005/2008/2011 de acordo com a matriz curricular de Matemática. O resultado apresentado servirá de base para a determinação de ações de ajuste a serem adotadas para a correção de qualquer desvio ou deficiência identificada no decorrer do curso superior, e antes da ocorrência da avaliação do ENADE, podendo servir posteriormente como ferramenta de uso recorrente na instituição.

A Etapa 2 refere-se à Elaboração dos itens e Aplicação do teste e consiste na elaboração de um teste de Matemática composto por 25 itens retirados de um banco de itens relativos às provas do ENADE/2005/2008/2011. Os itens foram escolhidos por uma equipe de professores de Matemática, do Departamento de Matemática e Física do CESC/UEMA. O teste foi aplicado para 45 alunos do 6º período do Curso Licenciatura em Matemática e teve duração de três horas, apresentou questões de múltipla escolha, relativas a conteúdo específico da área de Matemática.

Atendendo às exigências do ENADE 2011, no componente específico da área de Matemática, utilizou como referência o perfil de um profissional capaz de:

- i) conceber a Matemática como um corpo de conhecimentos rigoroso, formal e dedutivo, produto da atividade humana, historicamente construído;
- ii) dominar os conhecimentos matemáticos e compreender o seu uso em diferentes contextos interdisciplinares;
- iii) analisar criticamente a contribuição do conhecimento matemático na formação de indivíduos e no exercício da cidadania;
- iv) identificar, formular e solucionar problemas;
- v) valorizar a criatividade e a diversidade na elaboração de hipóteses, de proposições e na solução de problemas;
- vi) identificar concepções, valores e atitudes em relação à Matemática e seu ensino, visando à atuação crítica no desempenho profissional.

Para tanto, buscou-se investigar a formação de um profissional ético, competente e comprometido com a sociedade em que vive. Priorizando o domínio

de conhecimentos específicos com níveis diversificados de habilidades e competências para que evidenciem a compreensão de temas que transcendam ao seu ambiente próprio de formação e que sejam importantes para a realidade contemporânea.

Na Etapa 3, apresenta-se o resultado da aplicação e análise do teste. Nessa etapa foi criada uma planilha no Microsoft Excel, utilizando os dados do teste aplicado, em seguida a planilha foi salva como arquivo csv., para que fosse importada para o programa jMetrik, no qual foi realizada toda a análise estatística baseada nos fundamentos de TRI. Inicialmente foram considerados três aspectos: frequência de respostas, análise dos itens e análise da prova aplicada.

Para melhor compreensão do quadro abaixo, faz-se necessário recorrer à seguinte legenda:

- i) #A – número de alunos que marcaram a opção A;
- ii) #B – número de alunos que marcaram a opção B;
- iii) #C – número de alunos que marcaram a opção C;
- iv) #D – número de alunos que marcaram a opção D;
- v) #E – número de alunos que marcaram a opção E;

Os símbolos %A, %B, %C, %D e %E, correspondem ao percentual de alunos que marcaram cada item respectivamente.

Quadro 3 – Representação da frequência das respostas dadas pelos aluno

Questões	#A	%A	#B	%B	#C	%C	#D	%D	#E	%E
q1	12	27%	5	11%	15	33%	10	22%	3	7%
q2	6	13%	5	11%	12	27%	18	40%	4	9%
q3	5	11%	6	13%	9	20%	10	22%	15	33%
q4	6	13%	18	40%	5	11%	4	9%	12	27%
q5	2	4%	13	29%	0	0%	25	56%	5	11%
q6	1	2%	2	4%	12	27%	4	9%	26	58%
q7	28	62%	3	7%	8	18%	4	9%	2	4%
q8	16	36%	4	9%	14	31%	5	11%	6	13%
q9	7	16%	7	16%	7	16%	12	27%	12	27%
q10	6	13%	1	2%	23	51%	11	24%	4	9%
q11	4	9%	2	4%	5	11%	8	18%	26	58%
q12	7	16%	7	16%	9	20%	9	20%	13	29%
q13	8	18%	6	13%	8	18%	13	29%	10	22%
q14	7	16%	6	13%	9	20%	12	27%	11	24%
q15	11	24%	9	20%	9	20%	5	11%	11	24%
q16	3	7%	5	11%	7	16%	6	13%	24	53%
q17	15	33%	13	29%	3	7%	10	22%	4	9%
q18	6	13%	10	22%	14	31%	10	22%	5	11%
q19	7	16%	15	33%	12	27%	10	22%	1	2%
q20	20	44%	6	13%	5	11%	12	27%	2	4%
q21	3	7%	16	36%	11	24%	11	24%	4	9%
q22	10	22%	10	22%	9	20%	7	16%	9	20%
q23	4	9%	8	18%	12	27%	28	62%	9	20%
q24	9	20%	8	18%	9	20%	6	13%	13	29%
q25	7	16%	7	16%	16	36%	10	22%	5	11%

Fonte: A autora.

De acordo com a tabela acima se pôde perceber que as respostas dadas pelos alunos aos itens, foram bastante diversificadas, o que não possibilita para avaliar o grupo como um todo. Entretanto, destacam-se os itens: 5, 6, 10 e 11, que estão relacionados, respectivamente, aos conteúdos: solução de problemas do 1º grau envolvendo figuras espaciais; problemas envolvendo porcentagem; problemas envolvendo equações do 2º grau e construções geométricas. Tais itens apresentaram maior concentração de respostas em uma determinada alternativa.

Todavia, as informações obtidas não são suficientes para afirmar que os referidos alunos foram bem ou mal no teste. Desse modo, recorre-se à análise dos itens, o que garantirá uma informação mais precisa sobre a prova aplicada.

4.3 Análise resumida dos itens

A análise dos itens é fundamental para avaliar a qualidade dos testes desenvolvidos através de duas estatísticas: a dificuldade e a discriminação item.

Conforme Meyer (2013), essas estatísticas permitem identificar itens problemáticos tais como aqueles que são muito fáceis ou muito difíceis para os alunos e os itens que não estão relacionados à pontuação total. Como tal, segundo Hambleton e Swaminathan (1985), as pontuações de uma população heterogênea, no que diz respeito à característica medida, serão mais confiáveis que as pontuações a partir de uma população que é homogênea.

Assim, abaixo se descreve o resumo da análise feita pelo *Software* jMetrik, em que se destacam: a dificuldade do item, o índice de discriminação, Número de Respostas Erradas (N.R.E) e o Número de Respostas Corretas (N.R.C), bem como, em destaque, as alternativas corretas de cada item.

Quadro 4 – Resumo das estatísticas de itens do exame

ITEM	Dificuld.Item	Indc. Discrim.	N. R. E	N. R. C	A	B	C	D	E
1	0,33	0,02	30	15	12	5	15	10	3
2	0,11	0,10	40	5	6	5	12	18	4
3	0,20	-0,03	36	9	5	6	9	10	15
4	0,27	-0,35	33	12	6	18	5	4	12
5	0,56	0,01	20	25	2	13	0	25	5
6	0,58	-0,02	19	26	1	2	12	4	26
7	0,09	0,09	41	4	28	3	8	4	2
8	0,36	0,12	29	16	16	4	14	5	6
9	0,27	-0,05	33	12	7	7	7	12	12
10	0,51	-0,10	22	23	6	1	23	11	4
11	0,58	0,00	19	26	4	2	5	8	26
12	0,20	0,38	36	9	7	7	9	9	13
13	0,18	-0,19	37	8	8	6	8	13	10
14	0,20	-0,11	36	9	7	6	9	12	11
15	0,20	0,20	36	9	11	9	9	5	11
16	0,11	0,03	40	5	3	5	7	6	24
17	0,09	0,16	41	4	15	13	3	10	4
18	0,22	-0,16	35	10	6	10	14	10	5
19	0,27	0,04	33	12	7	15	12	10	1
20	0,27	0,10	33	12	20	6	5	12	2
21	0,24	-0,18	34	11	3	16	11	11	4
22	0,22	0,30	35	10	10	10	9	7	9
23	0,27	-0,15	33	12	4	8	12	12	9
24	0,13	0,22	39	6	9	8	9	6	13
25	0,11	-0,14	40	5	7	7	16	10	5

Fonte: A autora.

Para realizar uma análise de forma mais precisa, com base em Yu (2006), recorreu-se aos seguintes questionamentos, seguindo das suas respectivas respostas:

- i) Qual(is) item (s) você pode remover completamente a partir do exame? Por quê?

Ao explorar dificuldade do item, os itens 2, 7, 13, 16, 17, 24 e 25 parecem extremamente difíceis (Dificuldade do Item $<0,20$). Para a amostra analisada não existiram itens considerados muito fáceis (i. e. Dificuldade do Item $> 0,90$).

Ao explorar discriminação do item, os itens 1, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 16, 18, 20, 21, 23 e 25, parecem indicar pobres itens discriminativos (valores próximos ou inferiores a zero). Isso indica que os estudantes que obtiveram resultados ruins no exame como um todo, foram melhores nessas questões do que os alunos que, no geral, saíram-se bem. Infelizmente, no geral, o resultado da maioria dos alunos não foi significativo, resultando em muitos itens pouco discriminativos.

Portanto, combinando os resultados dessas duas investigações, os itens 3, 4, 6, 9, 10, 14, 18, 21 e 23, talvez pudessem ser considerados os itens mais confusos no teste, tais itens, parecem ser os melhores itens a serem excluídos deste exame.

- ii) Quais distratores você poderia revisar? Por quê?

Os itens 5, 6, 8, 10 e 11 parecem possuir distratores selecionados por poucos ou nenhum dos alunos, e os itens 7, 16, 17 e 25 parecem possuir distratores selecionados por tantos ou mais alunos do que a resposta correta. Por conseguinte, estes itens distratores devem ser removidos ou revistos, especialmente se os desviadores são a fonte de confusão dos alunos (em vez do produto em si).

- iii) Quais os itens que estão funcionando bem?

Itens que parecem estar funcionando bem (ou seja, boa dificuldade, discriminação e distratores) são itens, 2, 5, 8, 12, 20, 22, e 24 e, possivelmente, o 15 dependendo da finalidade do exame.

4.4 Análise da Prova

A interpretação da distribuição das respostas dos estudantes em cada questão da prova foi feita inicialmente a partir dos seguintes parâmetros descritivos de cada questão:

- i) índice de facilidade – proporção de participantes que responderam ao item corretamente;
- ii) correlação ponto bisserial entre a resposta correta no item e a pontuação total na prova;
- iii) a média do total de acertos dos participantes que acertaram um determinado item.

Os resultados revelaram que a prova apresenta questões difíceis (com índices de facilidade inferiores a 20% de acertos – questões: 2, 7, 13, 16, 17, 24 e 25). De acordo com a análise realizada, a prova não apresentou questões fáceis, pois nenhum índice de acertos foi superior a 90%. As questões difíceis envolvem conceitos como problemas envolvendo operações fundamentais, Limites e Derivadas Parciais.

Os resultados mostram as dificuldades apresentadas pelos estudantes, na sua maioria de conceitos básicos de Matemática, havendo assim, uma necessidade de retomada desses conceitos antes da realização do ENADE.

4.5 Interface do aplicador e modelagem do Sistema

A proposta deste trabalho baseia-se em uma modelagem para o processo de simulação de avaliação genérica, para os alunos do 6º período do Curso Licenciatura em Matemática do CESC/UEMA.

A seguir será apresentada e detalhada a proposta de modelagem do sistema identificando os atores deste processo: Alunos e Professores, bem como os diagramas de forma gráfica. Através da Linguagem de Modelagem Unificada (UML). Segundo Deitel (2003), a UML é um tipo de representação gráfica muito utilizado para modelagem de Sistemas Orientados a Objetos (OO), tornando-se uma das representações mais comuns.

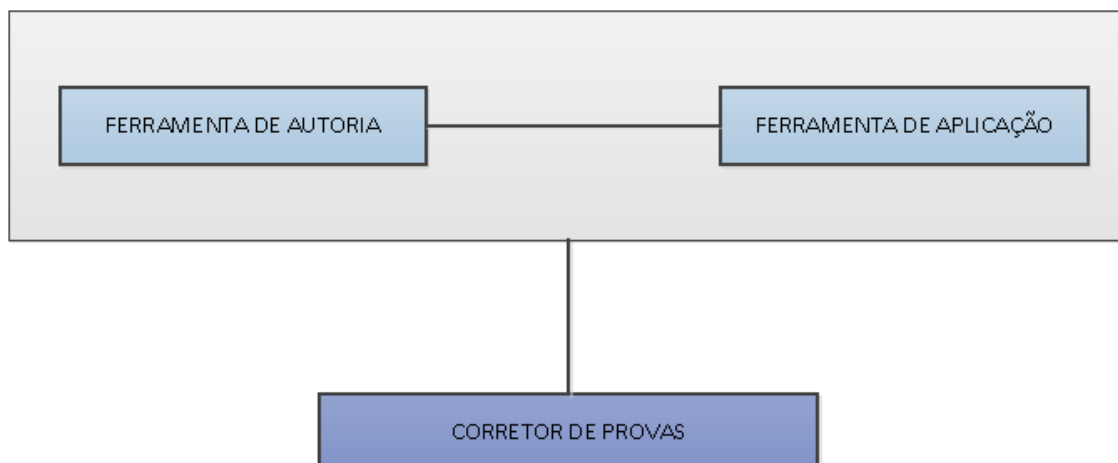
De acordo com Booch et al. (2005), a UML é uma linguagem gráfica para visualizar, especificar, construir e documentar os artefatos de um sistema de *software*. Por meio de seus diagramas, é possível representar sistema de *software* sob diversas perspectivas de visualização, diminuindo a complexidade no entendimento da construção de um sistema, facilitando a comunicação de todas as pessoas envolvidas no processo, além de apresentar um vocabulário de fácil compreensão.

Na primeira etapa, o sistema será preparado e adaptado para atender à realidade do curso em questão, em seguida, dar-se-á a preparação dos itens da prova pelos professores de Matemática, do Departamento de Matemática e Física do CESC/UEMA. Essa etapa será aprimorada ao longo do tempo gerando uma base de dados de questões para fomentar a diversidade quando da ocorrência de novas avaliações, seguindo pela validação dos itens pelo coordenador de forma a garantir a qualidade, coerência e coesão destes.

Ao elaborar os itens que farão parte da base de dados do Sistema, serão criado pelos professores além da chave de correção para cada item, sugestões de leituras como (artigos, objetos de aprendizagem, softwares específicos, etc.), caso o aluno dê uma resposta errada a um determinado item. Possibilitando, assim, um estudo mais aprofundado e compreensão significativa do conteúdo abordado. Contribuindo, dessa forma, para a construção do conhecimento. Em seguida, a composição da prova dar-se-á de forma aleatória e posteriormente será aplicada e corrigida de forma automática.

A figura abaixo ilustra a Arquitetura do Sistema, que é composta de três módulos. O primeiro que é a ferramenta de autoria de provas, acessada pelo ator autor, com seus casos de uso ilustrados na Figura 2. O segundo é a ferramenta de aplicação de provas, através da qual o respondente terá acesso às provas do sistema. O último é o corretor de provas.

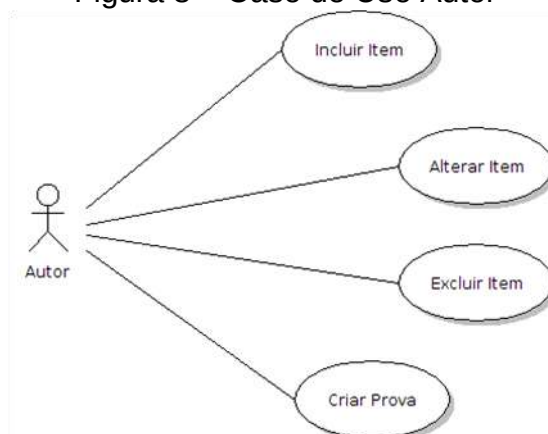
Figura 7 – Arquitetura Geral do Sistema



Fonte: A autora.

Para o primeiro módulo do sistema, a ferramenta de autoria, define-se um ator que lança mão desse módulo para alimentar a base de itens, é chamado de Autor. Ele é responsável por alimentar o Banco de Itens e criar as provas.

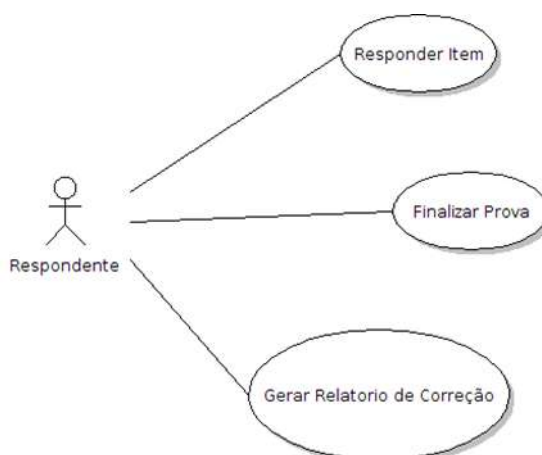
Figura 8 – Caso de Uso Autor



Fonte: A autora.

O ator que define o papel dos usuários no segundo módulo é o Respondente, ele é quem consome as provas e é avaliado. É para esse indivíduo que o relatório será gerado. O sistema prevê um relatório de desempenho, com sugestões de estudo em áreas em que o sistema detectar as maiores dificuldades do respondente, para tanto, no momento da formulação do item na ferramenta de autoria, o autor deverá também informar as referências bibliográficas e *links* para documentos contextuais ao item.

Figura 9 – Caso de Uso Aluno



Fonte: A autora.

O aluno, ao interagir com a ferramenta, será capaz de responder uma prova, e após o término o sistema analisará as respostas dadas pelos estudantes, obtendo assim *feedback* necessário que irá contribuir para um melhor aprendizado.

Na seção que se segue, será melhor detalhado o terceiro módulo do sistema, que implementa o corretor de provas TRI.

4.6 Uma proposta de corretor de provas TRI

O corretor de provas se constitui em uma ferramenta computacional que auxilia o professor a identificar as habilidades e os níveis de dificuldades dos alunos, bem como sugestões pedagógicas (em caso de respostas erradas), que propiciem a realização de estudos, com o objetivo de sanar dificuldades encontradas durante a prova.

4.6.1 Para entender melhor o problema vamos tomar o seguinte exemplo:

Uma escala de habilidades busca interpretar qualitativamente os valores obtidos pelos cálculos realizados por fórmulas da TRI, isto é, a escala é construída com a finalidade de saber quais os conceitos de uma determinada área os alunos sabem e quais eles ainda não aprenderam (VALLE, 1999). Por exemplo, na escala de habilidades do SAEB para conceitos de Matemática do 5º e 9º anos, de um aluno

que acertar questões de nível sete, presume-se que ele aprendeu ou já sabe os seguintes conceitos:

- i) Resolvem problemas com números naturais envolvendo diferentes significados da multiplicação e divisão, em situações combinatórias;
- ii) Reconhecem a conservação ou modificação de medidas dos lados, do perímetro, da área em ampliação e/ou modificação de medidas dos lados, do perímetro, da área em ampliação ou modificação de medidas dos lados, do perímetro, da área em ampliação e/ou redução de figuras poligonais usados malhas quadriculadas;
- iii) Identificam propriedades comuns e diferenças entre figuras bidimensionais pelo número de lados e tipos de ângulos;
- iv) Identificam as posições dos lados de quadriláteros (paralelismo);
- v) Resolvem problemas envolvendo utilização de divisão com resto diferente de zero, e estimam medidas de grandezas utilizando unidades de medidas convencionais;
- vi) Estabelecem relações entre unidades de medida de tempo;
- vii) Calculam resultado de uma divisão por meio de uma técnica operatória.

Os níveis da escala de habilidade em uma escala, para (Valle, 1999), são em geral organizados utilizando a escala (0:1), em que a média dos valores obtidos é zero e o desvio padrão (variabilidade) é igual a um. Por isso, alguns valores resultantes são menores que zero, e há a necessidade de se trabalhar com números que possuem várias casas decimais. Por exemplo, o aluno x pode ter seu traço latente igual a -1,52.

Para facilitar a criação de escalas, Valle (1999) explica que se podem usar outras técnicas para definir a métrica da escala.

Logo, por praticidade e também para facilitar o entendimento, é usual que se defina uma escala mais conveniente. Escolhe-se, por exemplo, um valor para a habilidade média de uma das populações, ou, então, se define que a escala deve variar apenas em um determinado range de valores. Uma vez definida a escala, faz-se uma transformação linear em todos os valores originais, para colocá-los na escala desejada. (VALLE, 2001, p. 75).

Após a escolha da métrica utilizada na escala, outro fator importante para o conhecimento das aptidões dos alunos é a identificação dos itens âncoras. Na concepção de Andrade et al. (2000) os níveis de habilidades de uma escala são caracterizados pelos itens âncora, ou seja, os itens âncoras de um teste são os responsáveis por localizar, dentre todos os itens, quais deles são os melhores para serem interpretados pelos educadores, a fim de identificar o nível a aptidão de aluno em uma área específica.

Um item é considerado âncora se e somente se ele satisfizer as seguintes condições:

Considere dois itens Y e Z (item âncora), sendo $Z > Y$.

- i) $P(U = 1 | \theta = Z) \geq 0,65$ e
- ii) $P(U = 1 | \theta = Y) \leq 0,50$ e
- iii) $P(U = 1 | \theta = Z) - P(U = 1 | \theta = Y) \geq 0,30$

Para um item ser âncora em um determinado nível âncora da escala, ele precisa ser respondido corretamente por uma grande proporção de indivíduos (pelo menos 65%) com esse nível de habilidade e por uma proporção menor de indivíduos (no máximo 50%) com nível de habilidade exatamente anterior. E a diferença entre a proporção de alunos com esses níveis de habilidade que acertam a esse item deve ser de no mínimo 30%.

Resumidamente, para um item ser considerado âncora é necessário que os indivíduos que possuem teta (θ) exatamente anterior ao do nível correspondente à dificuldade do item alcancem no máximo 30% de acertos neste item. Por exemplo, se um item com dificuldade 8 tiver 47% de acertos por alunos que possuem uma aptidão igual a teta (θ) igual a X, sendo X maior que Y, e tiver 38% de acertos por alunos com aptidão igual a Y, esse item não será classificado como item âncora, porque a sua capacidade de informar as características do educando não são satisfatórias, pois o número de alunos (com menos aptidões) que acertaram as questões faz do item possuir uma fraca capacidade de discriminar quais as aptidões do sujeito. Isso acontece porque o item foi acertado por várias pessoas de dois níveis diferentes, ou seja, os alunos com menos conhecimentos acertaram questões de nível mais alto. Esse fato ocorre toda vez que um novo item é desenvolvido e,

logo após sua aplicação, percebe-se que o item não informa corretamente qual a aptidão do aluno, a partir da interpretação dos resultados dos testes.

Salienta-se que um processo de avaliação deve contemplar uma perspectiva formativa, integrada com o planejamento e desenvolvimento de atividades pedagógicas. Nessa perspectiva, é imprescindível pensar numa avaliação continuada, que apresenta situações diversificadas e criativas e que envolva todo o processo de ensino e de aprendizagem.

Desse modo, a função do professor tem grande importância para a motivação do aluno, realizando *feedback*, para que esse se torne sujeito ativo e proativo.

Os processos de *feedback* e as orientações, como a orientação intrínseca/extrínseca, a orientação para a tarefa ou para o resultado e o nível de ansiedade exercem uma grande influência nas sensações de auto-estima, competência e autocontrole. (SAMULSKI, 2002, p. 117).

Considera-se que o *feedback* de um sistema de avaliação, é um processo de retroalimentação, que se caracteriza pela troca de informações entre os sujeitos, na medida que vão ocorrendo as interações e relações dentro do processo. Considera-se esse um dos pontos mais relevantes para o processo de avaliação, pois sua atuação é imediata e se for bem elaborado com certeza contribuirá de forma significativa na aprendizagem do aluno.

Nesta seção foi apresentada a proposta de desenvolvimento de uma ferramenta que tem como objetivo principal auxiliar no processo de avaliação da aprendizagem de estudantes do curso de Matemática Licenciatura do CESC/UEMA. Assim sendo, esta pesquisa, trata da criação de uma ferramenta que identifica as habilidades e proficiências dos discentes através da aplicação de testes e após essa identificação, o programa envia para o aluno materiais de estudo da *internet* compatíveis com as aptidões dos alunos.

Para tanto, a ferramenta propicia aos professores interfaces simples e interativas para criação e edição das provas, e ainda acesso a relatórios contendo informações sobre os níveis de aptidões dos discentes nos testes, bem como, dados estatísticos de porcentagem de erros e acertos das questões. Para os alunos, a ferramenta disponibiliza uma tela para realização das prova e mensagens de *feedback* que são enviadas pelo sistema. Com isso, podemos compreender que a

ferramenta contribui tanto na avaliação do ensino, como incentiva a aprendizagem de novos conceitos a partir dos *feedbacks* fornecidos.

5 CONSIDERAÇÕES FINAIS

Um dos principais problemas relacionados aos testes da ferramenta está relacionado à confecção de um banco de itens suficientemente grande para a aplicação exaustiva de provas. Consequentemente, os testes realizados até o momento são apenas em ambiente de desenvolvimento da aplicação. Apesar do cenário atual, os resultados são animadores e espera-se que em um prazo curto já se tenha a ferramenta suficientemente ajustada para que se possa colocá-la em um ambiente de produção.

Uma característica interessante dessa proposta é que, ao término da prova, é gerado um relatório automaticamente, no qual serão listadas as questões que foram respondidas de forma incorreta. Além disso, a ferramenta possibilitará ao aluno verificar seus erros e realizar estudos que propiciarão uma aprendizagem mais significativa, pois o próprio relatório já informa ao aluno indicações de estudo.

A ferramenta está sendo testada e serve para oferecer informações tanto para professor como para o aluno, simplificando o trabalho de correção além de prover um diagnóstico mais preciso a respeito dos alunos.

Com a análise dos dados realizada até o momento da realização desta pesquisa, conclui-se que a através da ferramenta é possível identificar problemas tanto estruturais como de ordem pedagógica e, sobretudo reavaliar as estratégias para o alcance dos objetivos e do perfil proposto, buscando nortear um trabalho que venha contribuir para melhor desempenho dos graduandos.

Ressalta-se que os resultados da prova realizada não foi animador, mas foi uma excelente oportunidade de análise para os professores do Departamento de Matemática e Física (re)pensarem o seu fazer pedagógico e, no coletivo, construir ações, metas, realizar intervenções pedagógicas para elevar o desempenho dos alunos, promovendo, assim, a equidade e a melhoria da qualidade da educação, bem como a necessidade de identificar com clareza o perfil dos alunos, compreendendo a realidade em que se encontram e os rumos que almejam atingir.

Para a criação dos casos de uso, recorreu-se aos fundamentos da UML, para especificar, visualizar e documentar modelos que representam as partes de sistema de software. No desenvolvimento do protótipo da ferramenta que tem como objetivo avaliar o rendimento escolar, foram utilizados para criar as páginas *web* Santos (2013), a tecnologia Linguagem de Marcação de Hipertexto – HTML (w3c, 2009) e a

linguagem de programação – Personal Home Page (PHP) –, Melo e Nascimento (2007). Entretanto, para armazenar as informações dos usuários, bem como as provas dos alunos, utilizou-se o banco de dados – Linguagem de Consulta Estruturada – MySQL (MySQL 2007).

Assim, a presente pesquisa preconiza uma necessidade contínua de investigar sobre avaliação, realçando a necessidade de técnicas que deve atender aos anseios pessoais e da qualidade do ensino oferecido à clientela em formação. E isso é um prenúncio de que muitas questões e reflexões podem ser suscitadas sobre a temática, a partir das constatações evidenciadas e das lacunas que por ventura forem percebidas neste estudo. Na certeza de que no percurso desta produção do conhecimento, sem grandes pretensões, poderá futuramente contribuir para a melhoria do processo de ensino e de aprendizagem, e acredita-se que muitos questionamentos surgirão a partir deste estudo como fonte e referência de pesquisa e de reflexões para outros pesquisadores no cenário educacional.

Parte deste trabalho foi aceito para apresentação no XXI Ciclo de Palestras Sobre Novas Tecnologias na Educação, realizado no período de 16 a 19 de julho de 2013 na cidade de Porto Alegre - RS, e foi publicado como artigo completo no periódico Revista Novas Tecnologias na Educação (RENOTE), da Universidade Federal do Rio Grande do Sul (UFRGS). Também foi aceito no 5º Simpósio Hipertexto e Tecnologias na Educação e 1º Colóquio Internacional de Educação com Tecnologias UFPE – Recife que será apresentado no período de 13 a 15 de novembro 2013. Pretende-se dar continuidade à pesquisa, submetendo-se outros artigos.

5.1 Contribuições deste trabalho

Diante dessas constatações, destaca-se que este estudo poderá contribuir com reflexões acerca da avaliação, percurso profissional do professor, da vivência e reflexão da prática pedagógica e também dos processos relativos à constituição de seus modos de ser professor, visto que os recursos tecnológicos da ferramenta foram desenvolvidos com a finalidade de auxiliar os discentes e docentes no processo de avaliação, incluindo as análises dos resultados obtidos na finalização do processo.

Dentre as contribuições, destacam-se:

- i) Automação do processo de cadastro e edição das questões e das provas;
- ii) Acesso a relatórios contendo informações sobre o rendimento coletivo e individual dos discentes nos testes.
- iii) Uma interface interativa para realização das provas;
- iv) Mensagens de *feedback* contendo referências literárias que estão de acordo com as dificuldades dos itens,
- v) Relatório com todos os resultados do discente nas avaliações.

A ferramenta disponibiliza funções básicas, para facilitar o acesso à informação dos testes, bem como o cadastro e gerenciamento de usuários no sistema. Tais como:

- i) Banco de dados contendo informações dos usuários, e ainda, todas as provas e resultados dos testandos;
- ii) Automação do cadastro e edição de todos os que utilizam o software.

5.2 Trabalhos Futuros

Como proposta para trabalhos futuros, entende-se que não foram esgotadas todas as possibilidades de estudo a respeito da temática desenvolvida nesta pesquisa. Ao se chegar às conclusões finais, deparou-se com novos questionamentos, vislumbrando, assim, um maior aprofundamento das questões.

Dentre os passos dados para maior aprofundamento, saliente-se que a autora tem direcionado atenção especial à situação que vem enfrentando o curso de Licenciatura em Matemática do CESC/UEMA. Pretende-se realizar um estudo de caso envolvendo todos os professores do Departamento de Matemática e Física e uma amostragem significativa dos alunos do curso em questão. Acredita-se que dessa forma será possível articular o perfil almejado para os discentes, utilizando uma ferramenta mais aperfeiçoada garantirá uma análise mais aprofundada da situação.

REFERÊNCIAS

- ABREU JÚNIOR, N. Sistema(s) de avaliação da educação superior brasileira. **Caderno CEDES**, Campinas, v. 29, n. 78, Ago. 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttextepid=S010132622009000200008&lng=en&rm=iso>. Acesso em: 01 jun. 2013.
- ABREU, M. C.; MASETTO, M. T. **O professor universitário em aula: prática e princípios teóricos**. São Paulo: Cortez, 1990.
- ALLEN, M. J.; YEN, W. M. **Introduction to measurement theory**. Illinois: Waveland Press, 2002.
- ANDRADE, Dalton Francisco. **Comparando o Desempenho de Grupos (populações) de Respondentes através da Teoria da Resposta ao Item**. PhD thesis, Tese (Doutorado). Departamento de Estatística e Matemática Aplicada. Universidade Federal do Ceará, 1999.
- _____.; TAVARES, Heliton Ribeiro; VALLE, Raquel da Cunha. **Teoria da Resposta ao Item: Conceitos e Aplicações**. SINAPE, 2000.
- ANDRADE, Josemberg Moura de; LAROS, Jacob Arie; GOUVEIA, Valdiney Veloso. **O uso da teoria de resposta ao item em avaliações educacionais: diretrizes para pesquisadores**. Aval. psicol., Porto Alegre, v. 9, n. 3, dez. 2010. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttextepid=S1677-04712010000300009&lng=pt&rm=iso>. Acesso em: 11 mar. 2013.
- ANDRICH, D. A Rating formulation for ordered response categories. **Psychometrika**, n. 43, p. 561-573, 1978.
- APÊNDICE – **Manual de utilização dos programas**. Disponível em: <http://www.inf.ufsc.br/~dandrade/Avaliacao_Educacional/manual.pdf>. Acesso em: 12 mai. 2013.
- ARAÚJO, E. A. C.; Andrade, D. F. de; Bortolotti, S. L. V. Teoria de Resposta ao Item. Disponível em: <<http://www.scielo.br/pdf/reeusp/v43nspe/a03v43ns.pdf>>. Acesso em: 25 jun. 2013.
- ARAÚJO, Gilda C. de; FERNANDES, Caroline F. R. Qualidade do ensino e avaliações em larga escala no Brasil: os desafios do processo e do sucesso educativo na garantia do direito à educação. **Revista Iberoamericana de Evaluación Educativa**, v. 2, n. 2, 2009.
- AVALIAÇÃO em Larga escala. Disponível em: <http://www.btdtd.ucb.br/tede/tde_busca/arquivo.php?codArquivo=1076>. Acesso em: 26 ago. 2013.
- BACKES, Dorimar Dal Bosco. **Avaliação do processo ensino aprendizagem: conceitos e Concepções**. Disponível em:

<http://www.nre.seed.pr.gov.br/cascavel/arquivos/File/Equipe%20Pedagogica/producao_dorimar.pdf>. Acesso em: 13 mar. 2013.

BIRNBAUM, A. Some Latent Traits Models and Their Use in Inferring an Examinee's Ability. In: **Statistical Theories of Mental Test Scores** [edited by F.M – Lord and M. R – Novick], 397 – 472. Addison – Wesley, Reading, MA, 1968.

BOCK, R. D.; ZIMOWSKI, M. F. Multiple group IRT. In: **Handbook of Modern Item response Theory**. W.J. van der Linder and R.K. Hambleton Eds. New York: Springer-Verlag, 1997.

_____. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. **Psychometrika**, v. 37, p. 29-51, 1972.

_____.; GIBBONS, R.; SCHILLING, S.; MURAKI, E. W.; WOOD, R. **Testfact 4 (Computer software)**. Lincolnwood, IL: Scientific Software International, 2003.

BONAMINO, Alicia; FRANCO, Creso. Avaliação e política educacional: o processo de institucionalização do Saeb. **Cadernos de Pesquisa**, São Paulo, n. 108, p. 101-132, nov. 1999.

BOOCH, Grady et al. UML: **Guia do Usuário**. Rio de Janeiro: Campus, 2000.

_____.; RUMBAUGH, James; JACOBSON, Ivair: **UML: Guia do Usuário**, Rio de Janeiro: Elsevier, 2005.

BRAGION, Maria de Lourdes Lima. **Um modelo de Teoria de Resposta ao Item para os dados do vestibular 2006-2 da UFLA**. Disponível em: <http://www.mch.ifsuldeminas.edu.br/~biblioteca/biblioteca_digital/Documentos/Doc_Maria%20de%20Lourdes%202.pdf>. Acesso em: 05 mar. 2013.

BRASIL. Ministério da Educação. Lei n. 10.861, de 14 de abril de 2004. Brasília: MEC, 2004. Disponível em: <http://www.inep.gov.br/download/superior/2004/Legislacao/LEI_n10861_14_4_04_SINAES>. Acesso em: 08 jun. 2013.

CARDOSO, RODRIGO FERRUGEM. AvalWeb – **Sistema interativo para gerência de questões e aplicação de avaliações na Web**. Disponível em: <http://www.ccei.urcamp.tc.br/disserta/disserta1_rodrigo.pdf>. Acesso em: 02 jun. 2013.

CASTRO, M.H.G. Sistemas de avaliação da educação no Brasil: avanços e novos desafios. **São Paulo em Perspectiva**, São Paulo, Fundação Seade, v. 23, n. 1, p. 5-18, jan./jun. 2009. Disponível em: <http://www.seade.gov.br/produtos/spp/v23n01/v23n01_01.pdf>. Acesso em: 06 mai. 2013.

COSTA, P. **Modelos de Resposta ao Item**. Dissertação (Mestrado). Universidade da Beira Interior, 2006;

CRESWELL, John W. **Projeto de Pesquisa: métodos qualitativo, quantitativo e misto** / Jonh W. Creswell; trad. Luciana de Oliveira da Rocha. 2. ed. Porto Alegre: Artmed, 2007.

DE AYALA, R. J.; SAVA-BOLESTA, M. Item parameter recovery for the nominal response model. **Applied Psychological Measurement**, 23, 3-19, 1999.

DEITEL, H. M.; DEITEL, P.J. **Java como Programar**. 4. ed. São Paulo: Makron Books, 2003.

DEMO, Pedro. **Teoria e Prática da Avaliação Qualitativa**. Temas do 2º Congresso sobre Avaliação na Educação. Curitiba, 2004.

DIAS SOBRINHO, José. **Avaliação da Educação Superior**. Petrópolis, RJ: Vozes, 2000.

_____. Educação superior sem fronteiras. Cenários da globalização: bem público, bem público global, comércio transnacional? **Avaliação: Revista da Rede de Avaliação Institucional da Educação Superior – RAIES**, Campinas, v. 9, n. 2, p. 9-29, jun. 2004.

EMILIANO, José Pereira. Javal - Modelo de Ambiente de Avaliação Remota Multiagente Baseada em Tutores Embarcados. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/4374/000455984.pdf?sequence=1>>. Acesso em 02 jun. 2013.

ENADE 2005. Disponível em: <<http://enadeies.inep.gov.br/enadeles/enadeResultado/>>. Acesso: 15 abr. 2013.

ENADE 2008. Disponível em: <<http://enadeies.inep.gov.br/enadeles/enadeResultado/>>. Acesso: 15 abr. 2013.

ENADE 2011. Disponível em: <<http://enadeies.inep.gov.br/enadeles/enadeResultado/>>. Acesso: 15 abr. 2013.

FERNANDES, Paula Gabriela de Medeiros. **Sistema Computadorizado de Avaliação Adaptativa em Larga Escala (SCAALE)**. Disponível em: <<http://monografias.cic.unb.br/dspace/bitstream/123456789/234/1/monografia.pdf>>. Acesso em: 12 jun. 2013.

FRANCISCO, Reinaldo. **Aplicação da Teoria da Resposta ao Item (TRI) no exame nacional de cursos (E.N.C) da Unicentro**. Disponível em: <<http://www.ppgmne.ufpr.br/arquivos/diss/121.pdf>>. Acesso em: 08 abr. 2013.

FRANCO, Creso. O SAEB – sistema de avaliação da educação básica: potencialidades, problemas e desafios. **Revista Brasileira de Educação**, 17, p. 127-33, 2001.

GARCÍA, Maria Izabel Barbero. **Desarrollos Recientes de los Modelos Psicométricos de la Teoría de Respuesta a los Ítems**. Disponível em: <<http://www.psicothema.com/pdf/242.pdf>>. Acesso em: 15 set. 2013.

GATTI, Bernardete A. Desenvolvimento de projetos de avaliação do sistema educacional no Estado de São Paulo. **Estudos em Avaliação Educacional**, 13, p. 19-26, 1996.

_____. **Avaliação de sistemas educacionais no Brasil**. Disponível em: <<http://sisifo.fpce.ul.pt/pdfs/Revista%209%20PT%20d1.pdf>>. Acesso em: 06 mai. 2013.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 1999.

_____. **Métodos e técnicas da pesquisa social**. São Paulo: Editora Atlas, 2010.

GONÇALVES, André Ricardo. **Aplicações de Técnicas de Aprendizagem de Máquina**. Disponível em: <<http://www2.dc.uel.br/nourau/document/?view=684>>. Acesso em: 05 ago. 2013.

GOTZMANN, Andrea. **Software Review- jMetrik item analysis (software application)**. Disponível em: <<http://www.rpajournal.com/dev/wp-content/uploads/2012/07/R3.pdf>>. Acesso em: 08 ago. 2013.

GUEWEHR, Katrine. **Teoria de resposta ao Item na avaliação de qualidade de vida dos idosos**. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/12638/000628234.pdf?sequence=1>>. Acesso em: 02 abr. 2013.

GUIA DE REFERÊNCIA CSS. Disponível em: <<http://www.w3c.br/pub/Materiais/PublicacoesW3C/guia-css-w3cbr.pdf>>. Acesso em: 13 jun. 2013.

HAMBLETON, R.; SWAMINATHAN, H. **Response Theory: principles and application**: Kluwer Publishing, 1985.

HEDGES, L. V.; VEVEA, J. L. A study of equating in NAEP. **Paper presented at The NAEP Validity Studies Panel**. Palo Alto: American Institutes for Research, 1997.

INEP. **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Disponível em: <www.inep.gov.br>. Acesso em: 09 jul. 2013.

ISSAC, E.; Keller, H. B. **Analysis of Numerical Methods**. New York: Wiley e Sons. 1966.

IAR – Instructional Assessment Resources. Disponível em: <<http://www.utexas.edu/academic/ctl/assessment/iar/students/report/itemanalysis.php>>. Acesso em: 05 ago. 2013.

JAVAL – **Modelo de Ambiente de Avaliação Remota Multiagente Baseada em Tutores Embarcados**. Disponível em:
<<http://www.lume.ufrgs.br/handle/10183/4374>>. Acesso em: 10 jun. 2013.

LEVINE, M.V.; DRASGOW, F. **Appropriateness measurement**: Validating studies and variable ability models. Dans D.J. Weiss (Éd.): *New horizons in testing – Latent trait test theory and computerized adaptive testing*. New York: Academic Press, 1983.

LIBÂNEO, José Carlos. **Didática**. São Paulo, Editora Cortez, 1991.

LORD, F. M. **A theory of test scores (No. 7)**. Psychometric Monograph, 1952.

_____. **Applications of Item Response Theory to Practical Testing Problems**. Hillsdale: Lawrence Erlbaum Associates, 1980.

_____, Norvick, M. R. **Statistical Theories of Mental Test Score**. Reading: Addison-Wesley, 1968.

MASTERS, G. N. A. Rasch model for partial credit scoring. **Psychometrika**, n. 47, p. 149-174, 1982.

MELO, Alexandre Altair de; NASCIMENTO, Mauricio G. F. **PHP Profissional**: Aprenda a desenvolver sistemas profissionais orientados a objetos com padrões de projeto. 1. ed. São Paulo: Novatec, 2007.

MEYER, J. Patrick. **Software Review**. Disponível em:
<<http://www.rpajournal.com/dev/wp-content/uploads/2012/07/R3.pdf>>. Acesso em: 05 ago. 2013.

_____. **Fair and Equitable Measurement of Student Learning in MOOCs**: An Introduction to Item Response Theory, Scale Linking, and Score Equating. Disponível em: <<http://www.rpajournal.com/dev/wp-content/uploads/2013/05/SF3.pdf>>. Acesso em: 13 set. 2013.

MISLEVY, R. J. **Bayes modal estimation in item response models**. *Psychometrika*, 51, p. 177–195, 1986.

MOREIRA JUNIOR, Fernando de Jesus. **Sistemática para implantação de testes adaptativos informatizados baseados na teoria da resposta ao item**. Disponível em:
<<http://repositorio.ufsc.br/bitstream/handle/123456789/95506/299657.pdf?sequence=1>>. Acesso em: 08 abr. 2013.

MURAKI, E. A generalized partial credit model: Application of an EM algorithm. **Applied Psychological Measurement**, n. 16, p. 159-176, 1992.

MYSQL. **Página oficial do banco de dados MySQL**. Disponível em:
<<http://www.mysql.com/>>. Acesso em: 05 mar. 2013.

PACHECO JÚNIOR, W.; PEREIRA, V. L. D. V.; PEREIRA FILHO, H. V. **Pesquisa científica sem tropeços**. São Paulo: Atlas, 2007;

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

PENIN, Sônia; MARTÍNEZ, Miguel. **Profissão docente**: pontos e contrapontos. São Paulo: Summus, 2009.

PEREIRA FILHO, Hippolito do Valle. **Pesquisa Científica Sem tropeços**: Abordagem Sistêmica. São Paulo: Atlas, 2007.

PEREIRA, Cleber Augusto. **Modelagem do sistema de avaliação de conhecimento, segundo parâmetros do Enade, aplicável aos cursos superiores de Graduação**: uma proposta quanto a forma de avaliação nas IES. Disponível em: <http://www.tedebc.ufma.br/tde_busca/arquivo.php?codArquivo=516>. Acesso em: 20 mar. 2013.

PEREIRA, Vinicius Ribeiro. **Métodos Alternativos no Critério Brasil para Construção de Indicadores Sócio-Econômico**: Teoria da Resposta ao Item. Disponível em: <http://www.maxwell.lambda.ele.puc-rio.br/5253/5253_4.PDF>. Acesso em: 20 ago. 2013.

PEREIRA, C. A.; LABIDI, S.; FONSECA L. C. C. **Modelagem do Sistema de Avaliação de Conhecimento Aplicável aos Cursos de Graduação, Segundo Parâmetros do ENADE**. Disponível em <<http://periodicos.unesc.net/index.php/sulcomp/article/view/286/294>>. Acesso em: 08 abr. 2013.

PERRENOUD, Philippe. **Ensinar**: agir na urgência, decidir na incerteza. Porto Alegre: Artes Médicas, 2002.

_____. MAGNE, B. C. **Construir**: as competências desde a escola. Porto Alegre: Artmed, 1999.

PINHEIRO, A.; SEN, P. K.; PINHEIRO, H. P. **A class of asymptotically normal degenerate quasi U-statistics**. Ann. Inst. of Stat. Math., pages1 – 18, 2010.

RABELLO, G. C. **A metodologia de equalização e o sistema nacional de avaliação da educação básica (Saeb)**, Technical report. Texto apresentado na série de Seminários SAEB 2001 – Estratégias para a Ação. 2001.

RAMOS, Edla Faust. O Papel da Avaliação Educacional nos Processos de Aprendizagem Autônomos e Cooperativos. In: LISIGEN, Irlan et al. **Formação do Engenheiro: Desafios da atuação docente, Tendências Curriculares e Questões da Educação Tecnológica**. Florianópolis: Editora da UFSC, 1999.

RAO, C. R. **Linear Statistical Inference and Its Applications**. New York: Wiley e Sons, 1973.

RASCH, G. **Probabilistic Models for Some Intelligence and Attainment Tests**. Institute for Educational Research, Copenhagen, Danish, 1960.

RELATÓRIO SÍNTESE DE MATEMÁTICA. Disponível em: <http://download.inep.gov.br/educacao_superior/enade/relatorio_sintese/2011/2011_rel_matematica.pdf>. Acesso em: 15 set. 2013.

RICHARDSON, R. H. **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 2010.

RODRIGUES, Malvina Thaís Pacheco. **Adesão ao tratamento da hipertensão arterial sistêmica**: desenvolvimento de um instrumento avaliativo com base na Teoria da Resposta ao Item (TRI). Disponível em: <www.uece.br/dsc/index.php/.../179-tese-malvina-t-p-rodrigues-2012>. Acesso em: 22 set. 2013.

SAMEJIMA, F. A. **Estimation of latent ability using a response pattern of graded scores**. Psychometric Monograph, n. 17, 1969.

SAMPAIO, Dulce Moreira. **Educação e a reconexão do Ser**: um caminho para a transformação humana e planetária. Petrópolis: Vozes, 2010.

SAMULSKI, D. **Psicologia do esporte**. Barueri: Manole, 2002.

SANTOS, Lanyllo Araújo dos. **Software para avaliar a aprendizagem escolar baseado nos conceitos da teoria de resposta ao item**. Codó, Universidade Federal do Maranhão, 2013.

SILVA, Jéssica Maria Afonso. **Modelo de Teoria de Resposta ao Item considerando efeitos de ansiedade**: como a ansiedade pode afetar o aluno frente ao teste. Disponível em: <<http://www.ufjf.br/cursoestatistica/files/2011/11/Monografia-J%C3%A9ssica-Maria.pdf>>. Acesso em: 02 ago. 2013.

SINAES. **Sistema Nacional de Avaliação da Educação Superior**: da concepção à regulamentação. 2. ed. Brasília: INEP, 2004.

TOIT, M. IRT from SSI: **BILOG-MG, MULTILOG, PARSCALE, TESTFACT**. Scientific Software International, 2003.

VALLE, Raquel da C. **Teoria da Resposta ao Item**. Dissertação de Mestrado. São Paulo: IME/USP, 1999.

_____. A construção e a interpretação das Escalas de Conhecimento – considerações gerais e uma visão do que vem sendo feito no SARESP. **Estudos em Avaliação Educacional**. Revista da Fundação Carlos Chagas, São Paulo, n. 23, p. 71-92, jan./jun. 2001.

VENDRAMINI, C. M. M.; SILVA, C. B.; CAZORLA, I. M.; BRITO, M. R. F. Análise da relação entre desempenho e atitudes em relação à Estatística. In: 14º SINAPE – Simpósio Nacional de Probabilidade e Estatística, 2000, Caxambú. 14º SINAPE –

Resumos. São Paulo: ABE – Associação Brasileira de Estatística, v. 2. p. 536-537. 2000.

VERAS, Jaclason Machado. **Modelagem para o software Virtual-TANEB baseado na Teoria da Resposta ao item para avaliar o rendimento dos alunos.** Disponível em:

<http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=ec_o_obra=169345>. Acesso em 22 mar. 2013.

VIANNA, Heraldo Marelim. Avaliação educacional e seus instrumentos: novos paradigmas. In: SOUSA, Eda C. B. Machado (Org.). **Curso de Especialização em Avaliação a Distância:** técnicas e instrumentos de avaliação. Brasília: UNB, 1997.

_____. **Fundamentos de um Programa de Avaliação Educacional.** Brasília: Liber Livro Editora, 2005.

WIEBUSCH, Eloisa Maria. **Avaliação em larga escala: uma possibilidade para a melhoria da aprendizagem.** Disponível em:

<<http://www.ucs.br/etc/conferencias/index.php/anpedsul/9anpedsul/paper/viewFile/1599/140>>. Acesso em: 05 ago. 2013.

_____. **Avaliação em Larga Escala: uma possibilidade para a melhoria da aprendizagem.** Disponível em:

<<http://www.ucs.br/etc/conferencias/index.php/anpedsul/9anpedsul/paper/viewFile/1599/140>>. Acesso em: 05 mai. 2013.

WRIGHT, B. D. **Sample-free test calibration and person measurement.**

Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N. J.: ETS – Educational Testing Service, 1968.

YANG, W.; HOUANG, R. T. **The effect of anchor length and equating method on the accuracy of test equating:** comparisons of linear and irt-based equating using anchor-item design. paper presented at the aera annual conference, division d., new york. zimowski, m. f., muraki, e., mislevy, r. j. e bock, r. d. (1996)., in 'BILOG-MG. Multiple-Group IRT analysis and test maintenance for binary items. Chicago: Scientific Software International.' 1996.

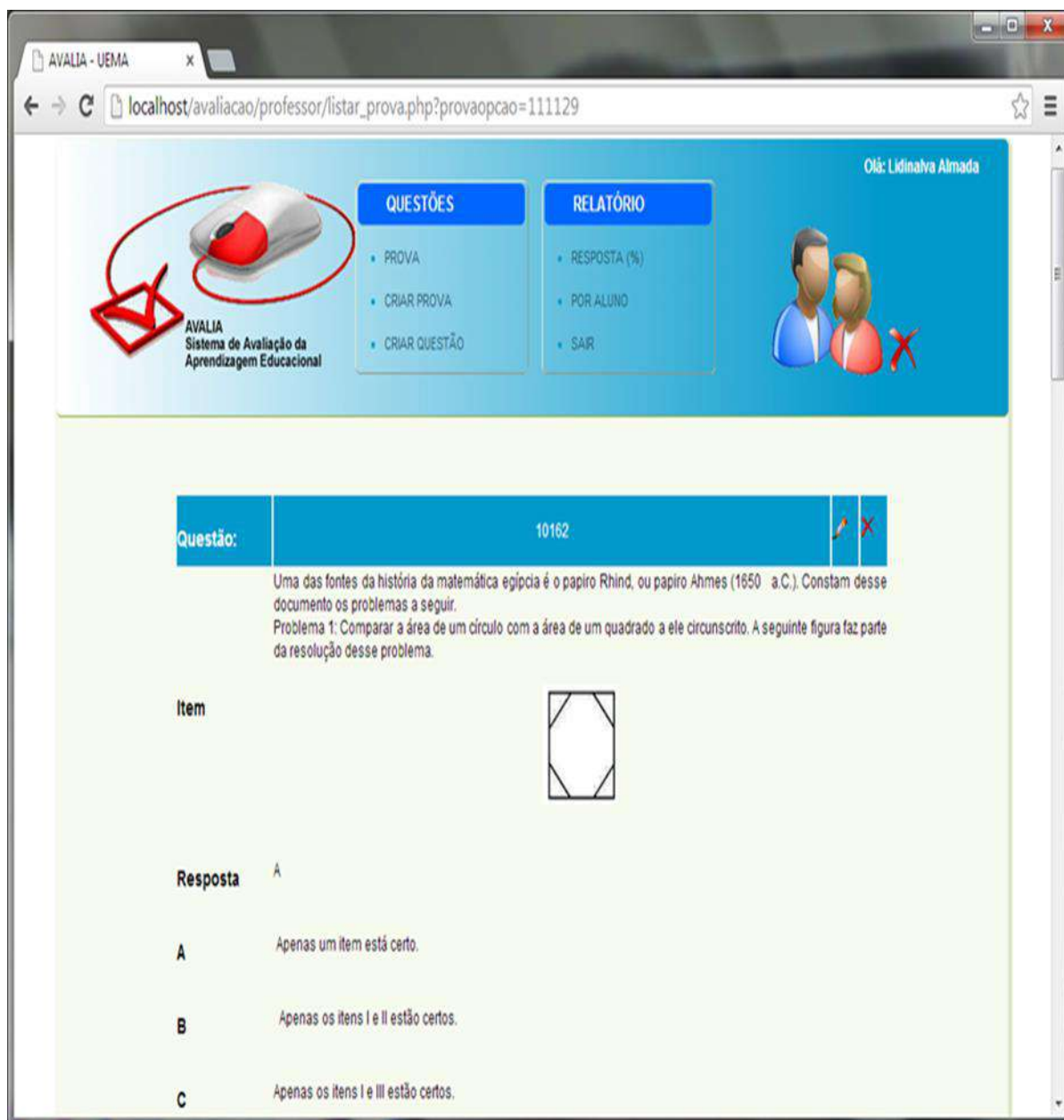
YU, A. (n.d) Using SAS for Item Analysis and Test Construction. Disponível em:

<<http://seamonkey.ed.asu.edu/~alex/teaching/assessment/alpha.html>>. Acesso em: 20 ago. 2013.

APÊNDICES

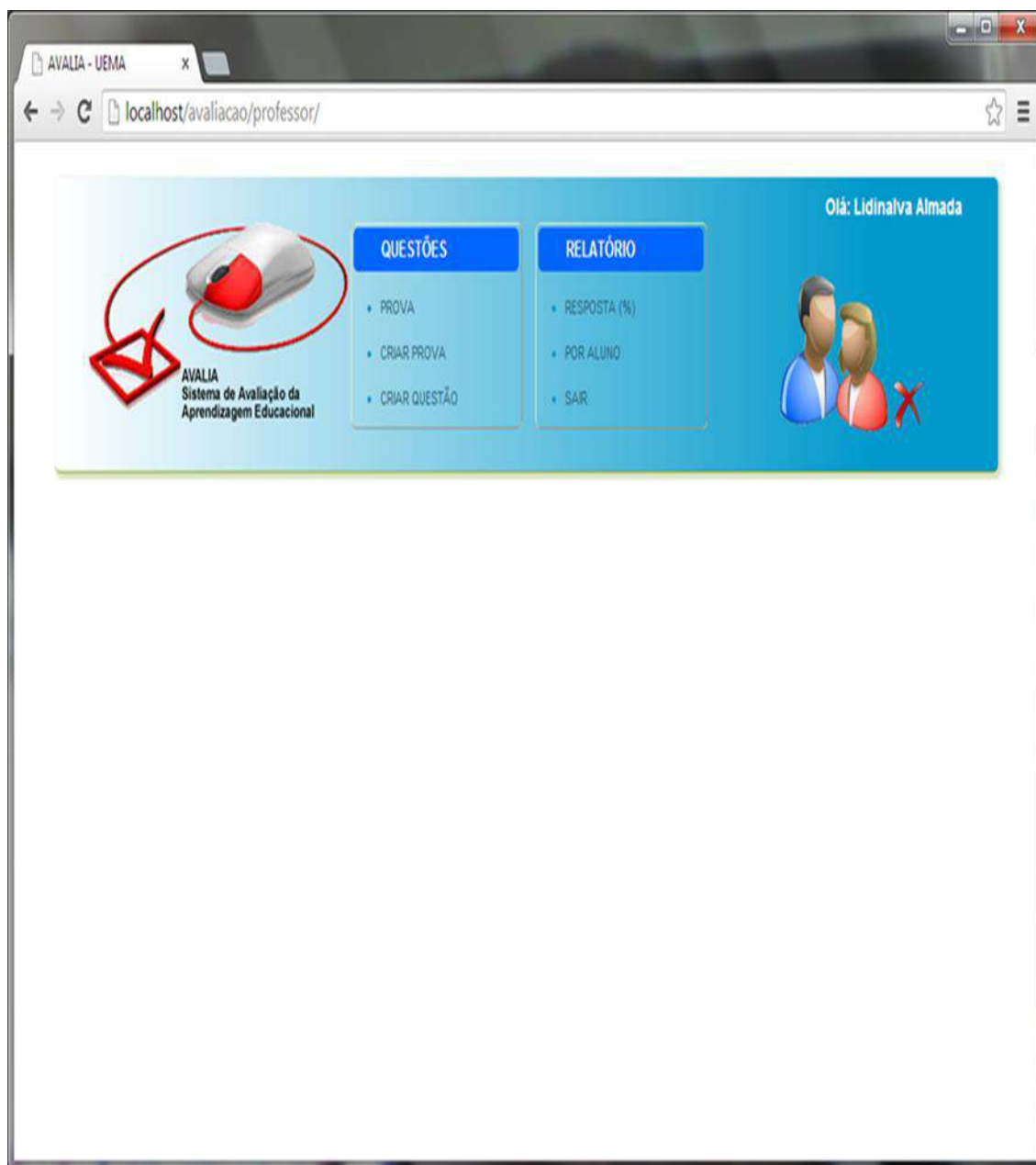
APÊNDICE-1

Figura 10: Interface para resolução das questões



Fonte: A autora

Figura 11: Área do professor na ferramenta



Fonte: A autora.

Figura 12: Formulário cadastro de Provas

AVALIA - UEMA

localhost/avaliacao/professor/form_cadastro_prova.php

Olá: Lidinalva Almada

QUESTÕES

- PROVA
- CRIAR PROVA
- CRIAR QUESTÃO

RELATÓRIO

- RESPOSTA (%)
- POR ALUNO
- SAIR

AVALIA
Sistema de Avaliação da
Aprendizagem Educacional

Início

Fim

Disciplina

Turma

Turno

Ano

Cancelar **Salvar**

Fonte: A autora.

Figura 13: Formulário de cadastro de questões

AVALIA Sistema de Avaliação da Aprendizagem Educacional

Olá: Lidinalva Almada

QUESTÕES

- PROVA
- CRIAR PROVA
- CRIAR QUESTÃO

RELATÓRIO

- RESPOSTA (%)
- FORALUNO
- SAIR

CADASTRAR NOVA QUESTÃO

Na aprendizagem da equação quadrática, a escola básica tende a trabalhar exclusivamente com a fórmula conhecida no Brasil como fórmula de Bhaskara. Entretanto, existem outras formulações desde a antiguidade, quando já se podiam identificar problemas e propostas de soluções para tais tipos de equação. Há mais de 4.000 anos, na Babilônia, adotavam-se procedimentos que hoje equivalem a expressar uma solução de como $2 + c$. Euclides (séc. I a.C.), no livro X de sua obra Os Elementos, já propunha uma resolução geométrica que permite resolver uma equação quadrática do tipo, utilizando exclusivamente compasso e régua não-graduada.

A respeito de uma proposta de ensino de resolução de equação quadrática com o enfoque em procedimentos historicamente construídos, assinale a opção correta.

Resposta

☐ A)

Fonte: A autora

Figura 14: Gabarito do discente



SIS-AVALIA UFMA

localhost/avaliacao/aluno/gabarito.php

Oi: Lidinalva Almada

AVALIA
Sistema de Avaliação da
Aprendizagem Educacional

MENU

- PROVA
- RESULTADOS
- RECOMENDAÇÕES

Mensagens (4)

Questão	Sua Resposta	Resposta Certa
10164	C	C
10164	C	C
10164	A	C
10162	B	A
10166	B	D
10164	A	C

Fonte: A autora.

Figura 15: Interface para resolução das questões

SIS-AVALIA UFMA

localhost/avaliacao/aluno/listar_prova.php?provaopcao=111129

Questão: Cod: 10165 Aluno(a) Lidinava Almada

Com o objetivo de chamar a atenção para o desperdício de água, um professor propôs a seguinte tarefa para seus alunos da 6.ª série do ensino fundamental:

Sabe-se que, em média, um banho de 15 minutos consome 136 L de água, o consumo de água de uma máquina de lavar roupas é de 75 L em uma lavagem completa e uma torneira pingando consome 46 L de água por dia. Considerando o número de banhos e o uso da máquina de lavar, compare a quantidade de água consumida por sua família durante uma semana com a quantidade de água que é desperdiçada por 2 torneiras pingando nesse período. Analise e comente os resultados.

No que se refere ao trabalho do aluno na resolução do problema proposto, assinale a opção incorreta.

- ☐ A) Elabora modelos matemáticos para resolver problemas.
- ☐ B) Analisa criticamente a situação-problema levando em conta
- ☐ C) Pode representar os resultados graficamente.
- ☐ D) Aciona estratégias de resolução de problemas.
- ☐ E) Examina consequências do uso de diferentes definições.

[Enviar Resposta](#)

Questão: Cod: 10166 Aluno(a) Lidinava Almada

Em uma classe da 6.ª série do ensino fundamental, o professor de matemática propôs aos alunos a descoberta de planificações para o cubo, que fossem diferentes daquelas trazidas tradicionalmente nos livros didáticos. Um grupo de alunos produziu a seguinte proposta de planificação.

Fonte: A autora.

Figura 16: Formulário de cadastro de Docentes e discentes

SIS-AVA UFMA

localhost/avaliacao/admin/form_cadastro_alu.php

AVALIA
Sistema de Avaliação da
Aprendizagem Educacional

CADASTRAR

- ALUNO
- PROFESSOR
- TURMA

BUSCA

- VER ALUNO
- VER PROFESSOR
- VER TURMA

Olá: Lidinalva Almada

Dados Pessoais

CPF *

044.591.493-97

Nome*

Data de Nascimento*

Senha*

Repita a senha*

Sexo

☒ Masculino ☐ Feminino

Estado Civil

Solteiro

Dados Residenciais

Endereço

Bairro

CEP*

Fonte: A autora

Figura 17: Edição de questões

Questão: 10166

Em uma classe da 6.ª série do ensino fundamental, o professor de matemática propôs aos alunos a descoberta de planificações para o cubo, que fossem diferentes daquelas trazidas tradicionalmente nos livros didáticos. Um grupo de alunos produziu a seguinte proposta de planificação.

Item

Ao tentar montar o cubo, o grupo descobriu que isso não era possível. Muitas justificativas foram dadas pelos participantes e estão listadas nas opções abaixo. Assinale aquela que tem fundamento matemático.

Resposta D

A Não se podem alinhar três quadrados.

B Tem de haver quatro quadrados alinhados, devendo estar os dois

C Quando três quadrados estão alinhados, não se pode mais ter os outros três também

D Cada ponto que corresponderá a um vértice deverá ser o encontro

E Tem de haver quatro quadrados alinhados, e não importa a

Sugestão

Fonte: A autora

Figura 18: Cadastro de Provas

SIS-AVA UFMA

localhost/avaliacao/admin/form_cadastro_turma.php

AVALIA
Sistema de Avaliação da
Aprendizagem Educacional

CADASTRAR

- ALUNO
- PROFESSOR
- TURMA

BUSCA

- VER ALUNO
- VER PROFESSOR
- VER TURMA

Olá: Lidinalva Almada

Turma 1A

Turno Vespertino

Ano 2013

Cancelar Salvar

Fonte: A autora.

Figura 19: Formulário de cadastro de questões

AVALIA - UEMA

localhost/avaliacao/professor/form_cadastro_item.php

Olá: Lidinalva Almada

QUESTÕES

- PROVA
- criar prova
- criar questão

RELATÓRIO

- resposta(n)
- por aluno
- sum

CADASTRAR NOVA QUESTÃO

Em uma classe da 6.ª série do ensino fundamental, o professor de matemática propôs aos alunos a descoberta de planificações para o cubo, que fossem diferentes daquelas trazidas tradicionalmente nos livros didáticos. Um grupo de alunos produziu a seguinte proposta de planificação.

Ao tentar montar o cubo, o grupo descobriu que isso não era possível. Muitas justificativas foram dadas pelos participantes e estão listadas nas opções abaixo. Assinale aquela que tem fundamento matemático.

Resposta

☐ A)

Fonte: A autora.

Figura 20: Interface para resolução das questões

☐ E) Examina consequências do uso de diferentes definições.

[Enviar Resposta](#)

Questão: Cod: 10100 Aluno(a): Lidiane Almada

Em uma classe da 6.ª série do ensino fundamental, o professor de matemática propôs aos alunos a descoberta de planificações para o cubo, que fossem diferentes daquelas trazidas tradicionalmente nos livros didáticos. Um grupo de alunos produziu a seguinte proposta de planificação.

Ao tentar montar o cubo, o grupo descobriu que isso não era possível. Muitas justificativas foram dadas pelos participantes e estão listadas nas opções abaixo. Assinale aquela que tem fundamento matemático.

- ☐ A) Não se podem alinhar três quadrados.
- ☐ B) Tem de haver quatro quadrados alinhados, devendo estar os dois
- ☐ C) Quando três quadrados estão alinhados, não se pode mais ter os outros três também
- ☐ D) Cada ponto que corresponderá a um vértice deverá ser o encontro
- ☐ E) Tem de haver quatro quadrados alinhados, e não importa a

Fonte: A autora

APÊNDICE- 2
ITEM ANALYSIS
testetril.PROVAUEMA
Outubro 12, 2013 15:16:19

Item	Option (Score)	Difficulty	Std. Dev.	Discrimin.
q1	Overall	0,3111	0,4682	0,0211
	A(0.0)	0,2667	0,4472	-0,2391
	B(0.0)	0,1111	0,3178	-0,2674
	C(1.0)	0,3111	0,4682	0,0211
	D(0.0)	0,2222	0,4204	-0,4342
	E(0.0)	0,0667	0,2523	0,0259
q2	Overall	0,1111	0,3178	0,1008
	A(0.0)	0,1333	0,3438	-0,2268
	B(1.0)	0,1111	0,3178	0,1008
	C(0.0)	0,2667	0,4472	-0,2833
	D(0.0)	0,4000	0,4954	-0,2609
	E(0.0)	0,0889	0,2878	-0,1402
q3	Overall	0,2000	0,4045	-0,0270
	A(0.0)	0,1111	0,3178	-0,0717
	B(0.0)	0,1333	0,3438	-0,3710
	C(1.0)	0,2000	0,4045	-0,0270
	D(0.0)	0,2222	0,4204	-0,1551
	E(0.0)	0,3333	0,4767	-0,2829
q4	Overall	0,2667	0,4472	-0,3480
	A(0.0)	0,1333	0,3438	-0,2853
	B(0.0)	0,4000	0,4954	-0,0963
	C(0.0)	0,1111	0,3178	-0,3302
	D(0.0)	0,0889	0,2878	0,2435
	E(1.0)	0,2667	0,4472	-0,3480
q5	Overall	0,5556	0,5025	0,0123
	A(0.0)	0,0444	0,2084	-0,1526
	B(0.0)	0,2889	0,4584	-0,2547
	C(0.0)	0,0000	0,0000	NaN
	D(1.0)	0,5556	0,5025	0,0123
	E(0.0)	0,1111	0,3178	-0,4222
q6	Overall	0,5778	0,4995	-0,0187
	A(0.0)	0,0222	0,1491	-0,1088
	B(0.0)	0,0444	0,2084	-0,3506
	C(0.0)	0,2667	0,4472	-0,2613
	D(0.0)	0,0889	0,2878	-0,2124
	E(1.0)	0,5778	0,4995	-0,0187

q7	Overall	0,0889	0,2878	0,0851
	A(0.0)	0,6222	0,4903	-0,2856
	B(0.0)	0,0667	0,2523	-0,1854
	C(0.0)	0,1778	0,3866	-0,1070
	D(1.0)	0,0889	0,2878	0,0851
	E(0.0)	0,0444	0,2084	-0,3018
q8	Overall	0,3556	0,4841	0,1215
	A(1.0)	0,3556	0,4841	0,1215
	B(0.0)	0,0889	0,2878	-0,1765
	C(0.0)	0,3111	0,4682	-0,3717
	D(0.0)	0,1111	0,3178	-0,1382
	E(0.0)	0,1333	0,3438	-0,3710
q9	Overall	0,2667	0,4472	-0,0537
	A(0.0)	0,1556	0,3665	-0,2474
	B(0.0)	0,1556	0,3665	-0,2474
	C(0.0)	0,1556	0,3665	-0,1069
	D(1.0)	0,2667	0,4472	-0,0537
	E(0.0)	0,2667	0,4472	-0,2833
q10	Overall	0,5111	0,5055	-0,0968
	A(0.0)	0,1333	0,3438	-0,1368
	B(0.0)	0,0222	0,1491	-0,1088
	C(1.0)	0,5111	0,5055	-0,0968
	D(0.0)	0,2444	0,4346	-0,4432
	E(0.0)	0,0889	0,2878	0,0084
q11	Overall	0,5778	0,4995	0,0035
	A(0.0)	0,0889	0,2878	-0,1402
	B(0.0)	0,0444	0,2084	-0,2028
	C(0.0)	0,1111	0,3178	-0,1382
	D(0.0)	0,1778	0,3866	-0,4162
	E(1.0)	0,5778	0,4995	0,0035
q12	Overall	0,2000	0,4045	0,3806
	A(0.0)	0,1556	0,3665	-0,2198
	B(0.0)	0,1556	0,3665	-0,5838
	C(0.0)	0,2000	0,4045	-0,2337
	D(1.0)	0,2000	0,4045	0,3806
	E(0.0)	0,2889	0,4584	-0,2331
q13	Overall	0,1778	0,3866	-0,1876
	A(0.0)	0,1778	0,3866	-0,2916
	B(0.0)	0,1333	0,3438	-0,1063
	C(1.0)	0,1778	0,3866	-0,1876
	D(0.0)	0,2889	0,4584	-0,2331
	E(0.0)	0,2222	0,4204	-0,1054

q14	Overall	0,2000	0,4045	-0,1065
	A(0.0)	0,1556	0,3665	-0,2198
	B(0.0)	0,1333	0,3438	-0,1368
	C(1.0)	0,2000	0,4045	-0,1065
	D(0.0)	0,2667	0,4472	-0,1713
	E(0.0)	0,2444	0,4346	-0,2906
q15	Overall	0,2000	0,4045	0,1978
	A(0.0)	0,2444	0,4346	-0,2906
	B(1.0)	0,2000	0,4045	0,1978
	C(0.0)	0,2000	0,4045	-0,3550
	D(0.0)	0,1111	0,3178	-0,0379
	E(0.0)	0,2444	0,4346	-0,3572
q16	Overall	0,1111	0,3178	0,0307
	A(0.0)	0,0667	0,2523	-0,1440
	B(1.0)	0,1111	0,3178	0,0307
	C(0.0)	0,1556	0,3665	-0,3017
	D(0.0)	0,1333	0,3438	-0,0127
	E(0.0)	0,5333	0,5045	-0,3253
q17	Overall	0,0889	0,2878	0,1634
	A(0.0)	0,3333	0,4767	-0,2829
	B(0.0)	0,2889	0,4584	-0,3185
	C(0.0)	0,0667	0,2523	0,1136
	D(0.0)	0,2222	0,4204	-0,3216
	E(1.0)	0,0889	0,2878	0,1634
q18	Overall	0,2222	0,4204	-0,1551
	A(0.0)	0,1333	0,3438	-0,3710
	B(1.0)	0,2222	0,4204	-0,1551
	C(0.0)	0,3111	0,4682	-0,1399
	D(0.0)	0,2222	0,4204	-0,2278
	E(0.0)	0,1111	0,3178	-0,0379
q19	Overall	0,2667	0,4472	0,0449
	A(0.0)	0,1556	0,3665	0,0106
	B(0.0)	0,3333	0,4767	-0,4595
	C(1.0)	0,2667	0,4472	0,0449
	D(0.0)	0,2222	0,4204	-0,3446
	E(0.0)	0,0222	0,1491	0,0356
q20	Overall	0,2667	0,4472	0,0958
	A(0.0)	0,4444	0,5025	-0,2452
	B(0.0)	0,1333	0,3438	-0,4269
	C(0.0)	0,1111	0,3178	-0,2356
	D(1.0)	0,2667	0,4472	0,0958
	E(0.0)	0,0444	0,2084	-0,1019

q21	Overall	0,2444	0,4346	-0,1755
	A(0.0)	0,0667	0,2523	-0,0599
	B(0.0)	0,3556	0,4841	-0,2143
	C(1.0)	0,2444	0,4346	-0,1755
	D(0.0)	0,2444	0,4346	-0,3352
	E(0.0)	0,0889	0,2878	-0,0293
q22	Overall	0,2222	0,4204	0,3034
	A(1.0)	0,2222	0,4204	0,3034
	B(0.0)	0,2222	0,4204	-0,2985
	C(0.0)	0,2000	0,4045	-0,2337
	D(0.0)	0,1556	0,3665	-0,3285
	E(0.0)	0,2000	0,4045	-0,3550
q23	Overall	0,2667	0,4472	-0,1482
	A(0.0)	0,0889	0,2878	-0,1402
	B(0.0)	0,1778	0,3866	-0,2916
	C(0.0)	0,2667	0,4472	-0,2613
	D(1.0)	0,2667	0,4472	-0,1482
	E(0.0)	0,2000	0,4045	-0,0803
q24	Overall	0,1333	0,3438	0,2177
	A(0.0)	0,2000	0,4045	-0,5160
	B(0.0)	0,1778	0,3866	-0,1341
	C(0.0)	0,2000	0,4045	-0,0803
	D(1.0)	0,1333	0,3438	0,2177
	E(0.0)	0,2889	0,4584	-0,2975
q25	Overall	0,1111	0,3178	-0,1382
	A(0.0)	0,1556	0,3665	-0,5095
	B(0.0)	0,1556	0,3665	0,0713
	C(0.0)	0,3556	0,4841	-0,0642
	D(0.0)	0,2222	0,4204	-0,2752
	E(1.0)	0,1111	0,3178	-0,1382

TEST LEVEL STATISTICS

```

=====
Number of Items = 25
Number of Examinees = 45
Min = 2,0000
Max = 13,0000
Mean = 6,5333
Median = 6,0000
Standard Deviation = 2,0827
Interquartile Range = 2,5000
Skewness = 0,4858
Kurtosis = 1,3319

```

KR21 = -0,1172

RELIABILITY ANALYSIS

Method	Estimate	95% Conf. Int.	SEM
Guttman's L2	0,1624	(-0,2350, 0,4786)	1,9062
Coefficient Alpha	0,0100	(-0,4597, 0,3837)	2,0723
Feldt-Gilmer	-0,0002	(-0,4748, 0,3774)	2,0830
Feldt-Brennan	0,0103	(-0,4593, 0,3839)	2,0954
Raju's Beta	0,0100	(-0,4597, 0,3837)	2,0723

Elapsed time: 0 secs, 256 msec

ANEXO

ANEXO

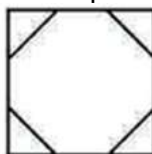
UNIVERSIDADE ESTADUAL DO MARANHÃO
CENTRO DE CIENCIAS TECNOLOGICAS
MESTRADO EM ENGENHARIA DA COMPUTAÇÃO E SISTEMAS

TESTE AVALIATIVO

Questão 1.

Uma das fontes da história da matemática egípcia é o papiro Rhind, ou papiro Ahmes (1650 a.C.). Constam desse documento os problemas a seguir.

Problema 1: Comparar a área de um círculo com a área de um quadrado a ele circunscrito. A seguinte figura faz parte da resolução desse problema.



Problema 2: “Exemplo de um corpo redondo de diâmetro 9. Qual é a área?”

A solução apresentada pelo escriba pode ser descrita como:

- remover $1/9$ do diâmetro; o restante é 8;
- multiplicar 8 por 8; perfaz 64. Portanto, a área é 64;

O procedimento do escriba permite calcular a área A de um círculo de diâmetro d aplicando a fórmula: $A = \left(\frac{8}{9}d\right)^2$

Com base nessas informações, julgue os itens a seguir.

- I A figura do problema 1 sugere aproximar a área de um círculo à área de um octógono.
- II O procedimento, no problema 2, fornece uma aproximação para π , por excesso, correta até a 2ª casa decimal.
- III De acordo com o procedimento, no problema 2, a área do círculo de diâmetro d é igual à de um quadrado de lado $8/9 d$.

Assinale a opção correta.

- A Apenas um item está certo.
- B Apenas os itens I e II estão certos.
- C Apenas os itens I e III estão certos.
- D Apenas os itens II e III estão certos.
- E Todos os itens estão certos.

Questão 2.

Na aprendizagem da equação quadrática, a escola básica tende a trabalhar exclusivamente com a fórmula conhecida no Brasil como fórmula de Bhaskara. Entretanto, existem outras formulações desde a antiguidade, quando já se podiam identificar problemas e propostas de soluções para tais tipos de equação. Há mais de 4.000 anos, na Babilônia, adotavam-se

procedimentos que hoje equivalem a expressar uma solução de $x^2 - bx = c$ com o $x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 + c}$. Euclides (séc. I a.C.), no livro X de sua obra Os Elementos, já propunha uma resolução geométrica que permite resolver uma equação quadrática do tipo $ax - x^2 = b$, utilizando exclusivamente compasso e régua não-graduada. A respeito de uma proposta de ensino de resolução de equação quadrática com o enfoque em procedimentos historicamente construídos, assinale a opção correta.

- A Tal proposta desvia a atenção da aprendizagem do foco central do conteúdo, fazendo que o aluno confunda as formulações, e, por consequência, não desenvolva competências na resolução de equações quadráticas.
- B É adequada a inserção dessa perspectiva, associada à manipulação de recorte e colagem pela complementação de quadrados, buscando sempre alternativas para as situações que esse procedimento não consegue resolver.
- C É mais adequado trabalhar o desenvolvimento da resolução de equações incompletas e, posteriormente, por meio da formulação de Bhaskara, manipular as equações completas, para somente no ensino médio ampliar tal conhecimento com o enfoque histórico.
- D É adequado utilizar tal proposta no ensino, uma vez que ela permite explicar a resolução de qualquer tipo de equação quadrática.
- E Tal proposta é inexecutável pelo tempo excessivo que exige do professor e por retardar a aprendizagem de alunos com dificuldades tanto em álgebra quanto em geometria.

Questão 3.

Não se pode negar que, embora bastante presentes em problemas envolvendo valores monetários e medidas, os números decimais constituem uma dificuldade no processo da aprendizagem matemática nas escolas. Uma das causas desse problema está na estrutura do currículo da matemática na escola básica.

Julgue os itens a seguir, acerca do ensino dos números decimais no currículo da educação básica.

- I Os números decimais representam uma expansão do sistema de numeração decimal enquanto base decimal e, por isso, seu conceito e representação no currículo precisam vir articulados à expansão da estrutura do sistema decimal.
- II O ensino dos números decimais deve preceder o ensino do sistema monetário, uma vez que o conhecimento dos decimais no currículo da educação básica é um pré-requisito para a aprendizagem desse conteúdo.
- III O currículo de matemática da escola básica deve propor, inicialmente, o ensino das frações com qualquer denominador, para então tratar das frações decimais como um caso específico, introduzindo, então, os números decimais.
- IV A ação do aluno em contextos de significado envolvendo valores monetários e medidas é fonte geradora de aprendizagem dos números decimais e, portanto, de ensino na escola, em um processo de resgate dos conhecimentos prévios dos alunos.

São reflexões apropriadas para a superação da problemática da baixa aprendizagem dos números decimais na escola apenas as contidas nos itens

- A I e II.
- B I e III.
- C I e IV.
- D II e III.
- E II, III e IV.

Questão 4.

Com o objetivo de chamar a atenção para o desperdício de água, um professor propôs a seguinte tarefa para seus alunos da 6.^a série do ensino fundamental:

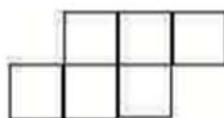
Sabe-se que, em média, um banho de 15 minutos consome 136 L de água, o consumo de água de uma máquina de lavar roupas é de 75 L em uma lavagem completa e uma torneira pingando consome 46 L de água por dia. Considerando o número de banhos e o uso da máquina de lavar, compare a quantidade de água consumida por sua família durante uma semana com a quantidade de água que é desperdiçada por 2 torneiras pingando nesse período. Analise e comente os resultados.

No que se refere ao trabalho do aluno na resolução do problema proposto, assinale a opção incorreta.

- A Elabora modelos matemáticos para resolver problemas.
- B Analisa criticamente a situação-problema levando em conta questões sociais.
- C Pode representar os resultados graficamente.
- D Aciona estratégias de resolução de problemas.
- E Examina consequências do uso de diferentes definições.

Questão 5.

Em uma classe da 6.^a série do ensino fundamental, o professor de matemática propôs aos alunos a descoberta de planificações para o cubo, que fossem diferentes daquelas trazidas tradicionalmente nos livros didáticos. Um grupo de alunos produziu a seguinte proposta de planificação.



Ao tentar montar o cubo, o grupo descobriu que isso não era possível. Muitas justificativas foram dadas pelos participantes e estão listadas nas opções abaixo. Assinale aquela que tem fundamento matemático.

- A Não se podem alinhar três quadrados.
- B Tem de haver quatro quadrados alinhados, devendo estar os dois quadrados restantes um de cada lado oposto dos quadrados alinhados.
- C Quando três quadrados estão alinhados, não se pode mais ter os outros três também alinhados.
- D Cada ponto que corresponderá a um vértice deverá ser o encontro de, no máximo, três segmentos, que serão as arestas do cubo.
- E Tem de haver quatro quadrados alinhados, e não importa a posição de justaposição dos outros dois quadrados.

Questão 6.

Julgue os itens a seguir, relativos ao ensino e à aprendizagem de porcentagens.

- I O ensino de porcentagem deve ter o contexto sociocultural como motivação de aprendizagem.
- II O primeiro contato dos estudantes com o cálculo percentual deve ocorrer quando se estudam juros compostos.
- III O ensino de frações centesimais e o de frações de quantidade devem ser articulados com o ensino de porcentagens.
- IV O conteúdo de porcentagens favorece um trabalho integrado entre diferentes blocos de conteúdos, tais como números, medidas, geometria e tratamento da informação.

Estão certos apenas os itens

- A I e II.
- B II e III.
- C III e IV.
- D I, II e III.
- E I, III e IV.

Questão 7.

Um aluno de 5.^a série, ao fazer a operação $63787 \div 3$ na resolução de um problema, foi considerado em “situação de dificuldade”, ao apresentar o seguinte registro:

$$\begin{array}{r}
 63787 \overline{) 3} \\
 00121 \quad 21222 \\
 \underline{-121} \quad \quad \quad 4 \\
 01 \quad 21262
 \end{array}$$

A análise do procedimento desse aluno revela que

- A Ele não sabe o algoritmo da divisão, o que indica problemas de aprendizagem oriundos das séries iniciais.
- B O procedimento aplicado não traz contribuições para o desenvolvimento matemático do aluno, uma vez que ele não poderá realizá-lo em outras situações matemáticas.
- C O aluno terá dificuldade de compreender os processos operatórios dos colegas e os feitos pelo professor ou apresentados no livro didático.
- D O aluno compreendeu tanto a estrutura do número quanto o conceito da operação de divisão.
- E Deverá ser incentivada a utilização de tal procedimento somente em produções individualizadas, como em atividades para casa.

Questão 8.

As potencialidades pedagógicas da história no ensino de matemática têm sido bastante discutidas. Entre as justificativas para o uso da história no ensino de matemática, inclui-se o fato de ela suscitar oportunidades para a investigação. Considerando essa justificativa, um professor propôs uma atividade a partir da informação histórica de que o famoso matemático Pierre Fermat [1601-1665], que se interessava por números primos, percebeu algumas relações entre números primos ímpares e quadrados perfeitos.

Para que os alunos também descobrissem essa relação, pediu que eles completassem a tabela a seguir, verificando quais números primos ímpares podem ser escritos como soma de dois quadrados perfeitos. Além disso, solicitou que observassem alguma propriedade comum a esses números.

3	5	7	11	13	17	19	23	29
	1+4			4+9	1+16			
não	sim	não	não	sim	sim			

A partir da atividade de investigação proposta pelo professor, analise as afirmações seguintes.

- I Todo número primo da forma $4n + 1$ pode ser escrito como a soma de dois quadrados perfeitos.
- II Todo número primo da forma $4n + 3$ pode ser escrito como a soma de dois quadrados perfeitos.
- III Todo número primo da forma $2n + 1$ pode ser escrito como a soma de dois quadrados perfeitos.

Está correto o que se afirma em

- A I, apenas.
- B II, apenas.
- C I e III, apenas.
- D II e III, apenas.
- E I, II e III.

Questão 9.

Na discussão relativa a funções exponenciais, um professor propôs a seguinte questão:

Para que valores não-nulos de k e m a função $f(x) = me^{kx}$ é uma função crescente?

Como estratégia de trabalho para que os alunos respondam à questão proposta, é adequado e suficiente o professor sugerir que os alunos:

- A Considerem $m = 1$ e $k = 1$, utilizem uma planilha eletrônica para calcular valores da função f em muitos pontos e comparem os valores obtidos.
- B Considerem $m = 1$ e $k = 1$, $m = -1$ e $k = 1$, esbocem os gráficos da função f e, em seguida, comparem esses dois gráficos.
- C Formem pequenos grupos, sendo que cada grupo deve esboçar o gráfico de uma das funções, para $m = 1, 2, 3, 4$ ou 5 , e comparem, em seguida, os gráficos encontrados.
- D Esbocem os gráficos das funções $y = e^x$ e $y = e^{-x}$ e analisem o que acontece com esses gráficos quando a variável e a função forem multiplicadas por constantes positivas ou negativas.
- E Construam uma tabela com os valores de f para x número inteiro variando de -5 a 5 , fixando $m = 1$ e $k = 1$ e, em seguida, comparem os valores encontrados.

Questão 10.

A professora Clara propôs a seus alunos que encontrassem a solução da seguinte equação do segundo grau:

$$x^2 - 1(2x + 3)(x - 1)$$

Pedro e João resolveram o exercício da seguinte maneira.

Resolução de Pedro:

$$\begin{array}{rcl} x^2 - 1 &= & (2x + 3)(x - 1) \\ x^2 - 1 &= & 2x^2 + x - 3 \\ 2 &= & x - x^2 \end{array}$$

Como 1 é solução dessa equação, então $S = \{1\}$

Resolução de João:

$$\begin{array}{rcl} x^2 - 1 &= & (2x + 3)(x - 1) \\ (x - 1)(x + 1) &= & (2x + 3)(x - 1) \\ x + 1 &= & 2x + 3 \\ x &= & 2 \end{array}$$

Portanto, $S = \{2\}$

Pedro e João perguntaram à professora por que encontraram soluções diferentes. A professora observou que outros alunos haviam apresentado soluções parecidas com as deles.

Entre as estratégias apresentadas nas opções a seguir, escolha a mais adequada a ser adotada por Clara visando à aprendizagem significativa por parte dos alunos.

- A Indicar individualmente, para cada aluno que apresentou uma resolução incorreta, onde está o erro e como corrigi-lo, a partir da estratégia inicial escolhida pelo aluno.
- B Resolver individualmente o exercício para cada aluno, usando a fórmula da resolução da equação do 2.º grau, mostrando que esse é o método que fornece a resposta correta.
- C Pedir a Pedro e João que apresentem à classe suas soluções para discussão e estimular os alunos a tentarem compreender onde está a falha nas soluções apresentadas e como devem fazer para corrigi-las.
- D Escrever a solução do exercício no quadro, usando a fórmula da resolução da equação do 2.º grau, para que os alunos percebam que esse é o método que fornece a resposta correta.
- E Pedir que cada um deles comunique à classe como resolveu o exercício e, em seguida, explicar no quadro para a turma onde está a falha na resolução de cada um e como eles devem fazer para corrigi-la.

Questão 11.

Observe a seguinte atividade de construções geométricas.

- Construir um triângulo ABC qualquer.
- Traçar a bissetriz do ângulo BAC e, em seguida, a bissetriz do ângulo ABC
- Marcar o ponto de encontro dessas duas bissetrizes.
- Traçar a bissetriz do ângulo ACB

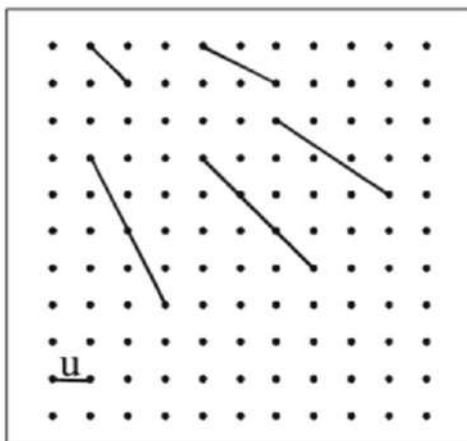
O que você observa?

Será que, se você recomençar a construção a partir de outro triângulo, chegará à mesma observação? O uso de um software de geometria dinâmica na execução dessa atividade e de outras similares

- A Pode mostrar que o estudo das construções com régua e compasso é desnecessário.
- B Dispensa a demonstração dos resultados encontrados pelos alunos.
- C Prejudica o desenvolvimento do raciocínio lógico-dedutivo.
- D Dificulta o desenvolvimento do pensamento geométrico.
- E Pode contribuir para a elaboração de conjecturas pelos alunos.

Questão 12.

A figura abaixo mostra alguns segmentos construídos em um geoplano por um estudante, de acordo com a orientação dada pela professora.



Acerca do uso do geoplano retangular nessa atividade, assinale a opção incorreta.

- A O geoplano auxilia na compreensão de que $\sqrt{a} + \sqrt{b} \neq \sqrt{a+b}$.
- B O geoplano auxilia na compreensão de que $\sqrt{ab} = \sqrt{a} \sqrt{b}$
- C O geoplano auxilia na representação geométrica de números irracionais da forma $\sqrt{a}$
- D O geoplano auxilia na obtenção da relação entre o comprimento de uma circunferência e seu diâmetro.
- E O geoplano auxilia na simplificação de expressões com irracionais algébricos, como, por exemplo, $\sqrt{20}\sqrt{5} = 3\sqrt{5}$

Questão 13.

Entre os procedimentos envolvidos na modelagem de uma situação-problema, estão sua tradução para a linguagem matemática e a resolução do problema, utilizando-se conhecimentos matemáticos. Nessa perspectiva, um professor propôs a seguinte situação-problema para seus alunos:

Escolha o nome para uma empresa que possa ser lido da mesma forma de qualquer um dos lados de uma porta de vidro transparente.

A solução desse problema pressupõe encontrar:

- A letras do alfabeto que sejam simétricas em relação a um ponto.
- B letras do alfabeto que tenham simetria em relação a um eixo horizontal.
- C letras do alfabeto que tenham simetria em relação a um eixo vertical.
- D palavras que sejam simétricas em relação a um ponto.
- E palavras que sejam simétricas em relação a um eixo horizontal.

Questão 14.

Considere o sistema de equações lineares $Ax = b$, com m equações e n incógnitas. Supondo que a solução do sistema homogêneo correspondente seja, única, avalie as afirmações a seguir.

- I. As colunas da matriz A são linearmente dependentes.
- II. O sistema de equações lineares $Ax = b$ tem infinitas soluções.

III. Se $m > n$, então a matriz A tem $m - n$ linhas que são combinações lineares de n linhas.
 IV. A quantidade de equações do sistema $Ax = b$ é maior ou igual à quantidade de incógnitas. São corretas apenas as afirmações

- A I e II.
- B II e III.
- C III e IV.
- D I, II e IV.
- E I, III e IV.

Questão 15.

Considere os elementos $\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ $\beta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ pertencentes ao grupo das permutações S_3 . Assinale a opção que representa $\alpha\beta$.

A $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$

B $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 1 \end{pmatrix}$

C $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$

D $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$

E $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$

Questão 16.

Sabe-se que, para todo número inteiro $n > 1$, tem-se $\frac{n \sqrt[n]{e}}{e} < \sqrt[n]{n!} < \frac{n \sqrt[n]{ne}}{e}$ Nesse

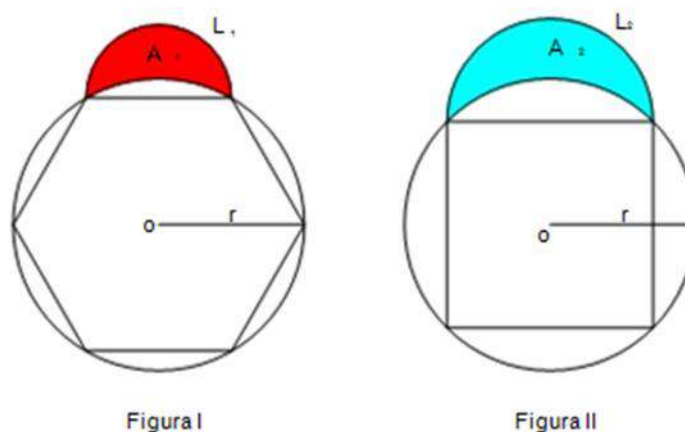
caso, se $\lim_{n \rightarrow \infty} \sqrt[n]{n!} = a$, então:

, $\lim_{n \rightarrow \infty} \frac{n \sqrt[n]{ne}}{e} = a$, então:

- A $a=0$
- B $a=1/e$
- C $a=1$
- D $a=e$
- E $a=+\infty$

Questão 17.

O matemático grego Hipócrates de Chios (470 a. C. - 410 a. C.) é conhecido como um excelente geômetra. Ele calculou a área de várias regiões do plano conhecidas como lúnulas, que são limitadas por arcos de circunferência, com centros e raios diferentes. As figuras I e II a seguir mostram, respectivamente, as lúnulas L_1 e L_2 , limitadas por um arco de circunferência de centro O e raio r e por semicircunferências cujos diâmetros são o lado de um hexágono regular e o lado de um quadrado inscritos na circunferência de raio r e centro O .



Considerando r um número racional, avalie as asserções a seguir.

A razão entre as áreas A_1 e A_2 das lúnulas L_1 e L_2 é um número racional.

PORQUE

A_1 e A_2 podem ser, respectivamente, representadas por $\pi \times q_1$ e $\pi \times q_2$, em que q_1 e q_2 são números racionais.

A respeito dessas asserções, assinale a opção correta.

- A As duas asserções são proposições verdadeiras, e a segunda é uma justificativa correta da primeira.
- B As duas asserções são proposições verdadeiras, mas a segunda não é uma justificativa da primeira.
- C A primeira asserção é uma proposição verdadeira, e a segunda, uma proposição falsa.
- D A primeira asserção é uma proposição falsa, e a segunda, uma proposição verdadeira.
- E Tanto a primeira quanto a segunda asserções são proposições falsas.

Questão 18.

O conjunto dos números complexos pode ser representado geometricamente no plano cartesiano de coordenadas xOy por meio da seguinte identificação:

$$z = x + iy \leftrightarrow P = (x, y)$$

Nesse contexto, analise as afirmações a seguir.

- I. As soluções da equação $z^4 = 1$ são vértices de um quadrado de lado 1.
- II. A representação geométrica dos números complexos z tais que $|z| = 1$ é uma circunferência com centro na origem e raio 1.
- III. A representação geométrica dos números complexos z tais que $\operatorname{Re}(z) + \operatorname{Im}(z) = 1$ é uma reta que tem coeficiente angular igual a $3\pi/4$ radianos.

É correto o que se afirma em:

- A I, apenas.
- B II, apenas.
- C I e III, apenas.
- D II e III, apenas.
- E I, II e III.

Questão 19.

Suponha que um instituto de pesquisa de opinião pública realizou um trabalho de modelagem matemática para mostrar a evolução das intenções de voto nas campanhas dos candidatos Paulo e Márcia a governador de um Estado, durante 36 quinzenas.

Os polinômios que representam, em porcentagem, a intenção dos votos dos eleitores de Paulo e Márcia na quinzena x são, respectivamente,

$$P(x) = -0,006x^2 + 0,8x + 14 \text{ e } M(x) = 0,004x^2 + 0,9x + 8,$$

em que $0 \leq x \leq 36$ representa a quinzena, $P(x)$ e $M(x)$ são dados em porcentagens.

De acordo com as pesquisas realizadas, a ordem de preferência nas intenções de voto em Paulo e Márcia sofreram alterações na quinzena

- A 6.
- B 12.
- C 20.
- D 22.
- E 30.

Questão 20.

Considere a função $f: R \rightarrow R$ definida por $f(x) = x^4 - 5x^2 = 4$, para cada $x \in R$. A área da região limitada pelo gráfico da função $y=f(x)$, o eixo Ox e as retas $x = 0$ e $x = 2$ é igual a

- A $16/15$ unidades de área.
- B $38/15$ unidades de área.
- C $44/15$ unidades de área.
- D $60/15$ unidades de área.
- E $76/15$ unidades de área.

Questão 21.

Considere $F: R^3 \rightarrow R$ uma função diferenciável e suponha que $F(x,y,z) = 0$ define implicitamente funções não nulas e diferenciáveis $z=f(x,y)$, $y=g(x,z)$ e $x=h(y,z)$. Nessa situação, analise as afirmações abaixo.

- I. $\frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x} \left(1 + \frac{1}{n}\right)^n$
- II. Se $F_z(x, y, z) \neq 0$, então $\frac{\partial z}{\partial x} = -\frac{F_x(x, y, z)}{F_z(x, y, z)}$.
- III. $\frac{\partial z}{\partial x} \frac{\partial x}{\partial y} \frac{\partial y}{\partial z} = 1$

É correto o que se afirma em:

- A II, apenas.
- B III, apenas.
- C I e II, apenas.
- D I e III, apenas.
- E I, II e III.

Questão 22.

Catedral Metropolitana de Brasília

A construção da Catedral, projeto do arquiteto Oscar Niemeyer, teve início em 12 de agosto de 1958, em plena construção da nova capital. Em 1959, mesmo antes da inauguração de Brasília (1960), a sua forma estrutural (pilares de concreto armado, na forma de um hiperbolóide de revolução) já estava pronta. O fechamento lateral entre os pilares só ocorreu em 1967, pouco antes de sua consagração, em 12 de outubro do mesmo ano, ocasião em que recebeu a imagem de Nossa Senhora Aparecida.

De 1969 a 1970, o complexo foi concluído com o espelho d'água ao redor da Catedral, o batistério e o campanário.

PORTO, C. E. Um estudo comparativo da forma estrutural de dois monumentos religiosos em Brasília: A Catedral e o Estupa Tibetano. Disponível em: <www.skyscraperlife.com/arquitetura-e-discussoes-urbanas/22122-obras-de-oscar-niemeyer.html>. Acesso em 30 ago. 2011.



Figura 1 - Catedral Metropolitana de Brasília.

Nesse contexto, considere na figura abaixo os elementos principais da hipérbole associada aos arcos hiperbólicos da Catedral Metropolitana de Brasília.

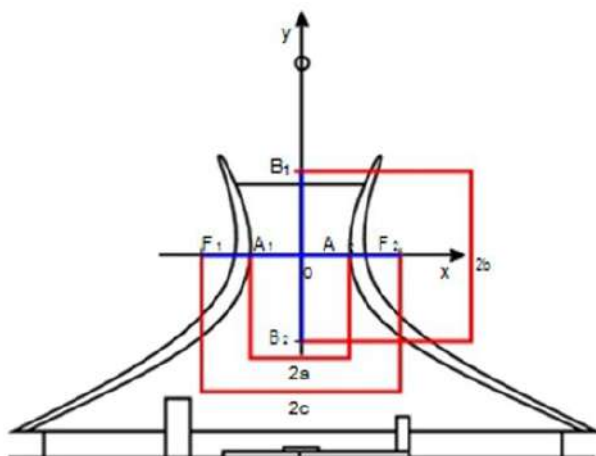


Figura II Corte esquemático da Catedral, representando os arcos hiperbólicos (Niemeyer, 1958, p. 14)

Supondo que o eixo real (ou eixo transversal) da hipérbole na figura II mede 30 m e que a distância focal mede 50 m, analise as seguintes asserções.

Se $F_1 = (-c, 0)$ é o foco da hipérbole, então a diretriz associada a ela é a reta $d_1: x + 9 = 0$.

PORQUE

A equação reduzida dessa hipérbole é $\frac{x^2}{225} - \frac{y^2}{400} = 1$

- A As duas asserções são proposições verdadeiras, e a segunda é uma justificativa correta da primeira.
- B As duas asserções são proposições verdadeiras, mas a segunda não é uma justificativa da primeira.
- C A primeira asserção é uma proposição verdadeira, e a segunda, uma proposição falsa.
- D A primeira asserção é uma proposição falsa, e a segunda, uma proposição verdadeira.
- E Tanto a primeira quanto a segunda asserções são proposições falsas.

Questão 23.

Considerando a , b e c pertencentes ao conjunto dos números naturais e representando por $a|b$ a relação “ a divide b ”, analise as proposições abaixo.

- I. Se $a|(b + c)$, então $a|b$ ou $a|c$.
 - II. Se $a|bc$ e $\text{mdc}(a, b) = 1$, então $a|c$.
 - III. Se a não é primo e $a|bc$, então $a|b$ ou $a|c$.
 - IV. Se $a|b$ e $\text{mdc}(b, c) = 1$, então $\text{mdc}(a, c) = 1$.
- É correto apenas o que se afirma em:

- A I.
- B II.
- C I e III.
- D II e IV.
- E III e IV.

Questão 24.

Os analistas financeiros de uma empresa chegaram a um modelo matemático que permite calcular a arrecadação mensal da empresa ao longo de 24 meses, por meio da função $A(x) = \frac{x^3}{3} - 11x^2 + 117x + 4$, em que $0 \leq x \leq 24$ é o tempo, em meses, e a arrecadação $A(x)$ é dada em milhões de reais.

A arrecadação da empresa começou a decrescer e, depois, retomou o crescimento, respectivamente, a partir dos meses

- A $x = 0$ e $x = 11$.
- B $x = 4$ e $x = 7$.
- C $x = 8$ e $x = 16$.
- D $x = 9$ e $x = 13$.
- E $x = 11$ e $x = 22$.

Questão 25.

Considere $u(x, y) = f(x - 4y) + g(x + 4y)$, em que f e g são funções reais quaisquer, deriváveis até a segunda ordem, com $U_{xx} \neq 0$ para todo x e y . Nesse caso, $\frac{U_{yy}}{U_{xx}}$ é igual a

- A - 16 .
- B - 8 .
- C 0.
- D 8.
- E 16.