

UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE ELETRICIDADE
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA DE ELETRICIDADE

ANA CAROLINA MELO DE OLIVEIRA FARIAS

**APLICAÇÃO DA TEORIA DE RESPOSTA AO ITEM NA
AVALIAÇÃO EDUCACIONAL E SEU PROCESSO DE
INFORMATIZAÇÃO**

São Luís

2013

ANA CAROLINA MELO DE OLIVEIRA FARIAS

**APLICAÇÃO DA TEORIA DE RESPOSTA AO ITEM NA
AVALIAÇÃO EDUCACIONAL E SEU PROCESSO DE
INFORMATIZAÇÃO**

Dissertação de Mestrado apresentada ao curso de Pós Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, como parte dos requisitos para a obtenção do título de Mestre em Engenharia de Eletricidade, na área de Ciência da Computação.

Orientador: Prof. Dr. Sofiane Labidi

São Luís

2013

Farias, Ana Carolina Melo de Oliveira.

Aplicação da teoria de resposta ao item na avaliação educacional e seu processo de informatização / Ana Carolina Melo de Oliveira Farias – São Luís, 2013.

78 f.

Impresso por computador (fotocópia).
Orientador: Sofiane Labidi.

Dissertação (Mestrado) – Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia de Eletricidade, 2013.

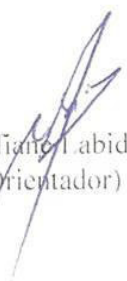
1. Avaliação educacional. 2. Sistema de Avaliação. 3. Teoria de Resposta ao Item. 4. Teorias de Medidas. I. Título.

CDU 621.3:373.3

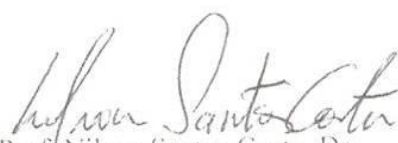
APLICAÇÃO DA TEORIA DE RESPOSTA AO ITEM NA AVALIAÇÃO EDUCACIONAL E SEU PROCESSO DE INFORMATIZAÇÃO

Ana Carolina Melo de Oliveira Farias

Dissertação aprovada em 04 de março de 2013.



Prof. Sofiane Labidi, Dr.,
(Orientador)



Prof. Nilson Santos Costa, Dr.,
(Membro da Banca Examinadora)



Profa. Maria do Perpetuo Socorro Soares Teixeira, Dra.,
(Membro da Banca Examinadora)



Prof. Dalton Francisco de Andrade, Dr.,
(Membro da Banca Examinadora)

*Aos meus pais que sempre se dedicaram a mim e
acreditaram no meu sucesso.*

AGRADECIMENTOS

Primeiramente devo agradecer a Deus pelas bênçãos derramadas na minha vida, que por meio de orações, me alimentam de fé e esperança não permitindo que eu desista dos meus sonhos.

Aos meus pais que souberam me conduzir e nunca pouparam esforços e dedicação na minha educação.

Aos meus irmãos, que são os meus melhores amigos e sempre tiveram paciência em me ouvir.

Ao meu esposo, que além de amigo é um grande companheiro.

Ao meu filho, que apesar da sua pouca idade sempre me conforta com palavras carinhosas nos momentos em que preciso.

Ao meu padrinho Raimundo Barroso, que sempre se mostrou disponível em ajudar-me e orientar-me.

A professora Socorro Teixeira, pelo esforço e dedicação oferecida durante toda a trajetória percorrida para o alcance desta conquista.

Ao professor Sofiane Labidi, por me receber e me orientar com todo seu carisma e compreensão, tornando possível a realização deste trabalho.

Ao professor Nilson Costa, pela ajuda, orientação e acompanhamento constante durante todo o trabalho.

Ao professor Dalton Andrade, que mesmo distante se mostrou presente e dedicou toda sua experiência e atenção mostrando-se solícito a todas as minhas dúvidas.

A todos os alunos e ex-alunos do LSI, em especial a Pedriana Pavão que me apoiou e me ajudou desde o início deste trabalho.

A todos que me ajudaram e participaram dessa conquista, aos meus amigos e a toda minha família, meus sinceros e eternos agradecimentos.

*É o tempo da travessia...
E se não ousarmos fazê-la...
Teremos ficado para sempre...
À margem de nós mesmos...*

(Fernando Pessoa)

RESUMO

O processo educacional existente atualmente utiliza o computador como ferramenta auxiliar no processo de ensino-aprendizagem. Essa ferramenta vem ganhando espaço em âmbito educacional, sendo assim adotada em algumas escolas. O processo de avaliação educacional busca formas de avaliar que garanta a segurança dos itens e a rapidez e agilidade nos resultados, e por isso já destaca a importância de inserir o computador ao seu processo. Este trabalho trata do uso do computador no processo de avaliação aplicado a alunos do 5º ano do ensino fundamental associado ao uso da Teoria de Resposta ao Item no processo de avaliação e obtenção de resultados. Dessa forma, foi proposto o desenvolvimento de um módulo de avaliação online que possibilite a realização de um teste no computador. Este mesmo teste foi aplicado no papel e caneta para que assim, como forma de experimento, seja possível com o uso da TRI e por meio de um recurso computacional, avaliar os resultados e posteriormente fazer um levantamento das vantagens e desvantagens de inserir o computador no processo de avaliação, e justificar os diferentes comportamentos apresentados pelos alunos que se submeteram as provas.

Palavras-Chave: Avaliação Educacional. Sistemas de Avaliação. Teoria de Resposta ao Item. Teorias de Medidas.

ABSTRACT

The existing educational process currently uses the computer as an auxiliary tool in the teaching-learning process. This tool has been growing up in the educational field, and this has already been adopted in some schools. The process of educational assessment seeks ways to evaluate to ensure the safety of items and speed and agility in the results, for this the importance of entering the computer to its process. This work deals with the use of computers in the assessment process used with students from 5th grade key associated with the use of Item Response Theory (IRT) in the evaluation process and results. Thus, we developed a web-based assessment module that fostered the development of a test on the computer. The same test was applied to paper and pen so that as a way to experiment is possible with the use of ITR, through software BILOG, evaluate the results and subsequently make a survey of the advantages and disadvantages of entering the computer in the process evaluation, and justify the different behaviors presented by students who submitted the evidence.

Keywords: Educational Assessment. Evaluation Systems. Item Theory of Response . Measurement Theory.

LISTA DE FIGURAS

Figura 1	- Curva Característica do Item.....	36
Figura 2	- BILOG – MG.....	46
Figura 3	- Modelagem do Módulo de Avaliação Proposto.....	48
Figura 4	- Modelagem BPMN do Sistema de Avaliação.....	50
Figura 5	- Diagrama de Caso de Uso “Fazer – Prova”.....	51
Figura 6	- Diagrama de Caso de Uso “Corrigir – Prova”.....	52
Figura 7	- Módulo de Avaliação – Tela “Cadastro”.....	56
Figura 8	- Módulo de Avaliação – “Fazer Login”.....	56
Figura 9	- Item 12 do Módulo de Avaliação aplicado no computador para meninas.....	57
Figura 10	- Item 12 do Módulo de Avaliação aplicado no computador para meninos.....	57
Figura 11	- Item 12 do Módulo de Avaliação aplicado no papel para meninos e meninas.....	58
Figura 12	- Relatório Final de Avaliação.....	59
Figura 13	- Representação Gráfica da aplicação da prova “Dois grupos fazendo uma única prova”.....	60
Figura 14	- Item 02 do Módulo de Avaliação aplicado no papel para meninos e meninas.....	62
Figura 15	- Item 04 do Módulo de Avaliação aplicado no papel para meninos e meninas.....	63
Figura 16	- Item 22 do Módulo de Avaliação aplicado no computador para meninos.....	64
Figura 17	- Item 24 do Módulo de Avaliação aplicado no computador para meninos.....	64

Figura 18	-	Aplicação do teste em laboratório para meninos e meninas....	67
Figura 19	-	Item 02 do Módulo de Avaliação aplicado no computador para meninas.....	68
Figura 20	-	Item 04 do Módulo de Avaliação aplicado no computador para meninas.....	68
Figura 21	-	Gráfico Comparativo entre a prova escrita e no computador...	71

LISTA DE TABELAS

Tabela 1	- Modelos Matemáticos.....	34
Tabela 2	- Calibração na Escala 0,1 da prova escrita.....	61
Tabela 3	- Calibração na Escala 0,1 da prova no computador.....	66
Tabela 4	- Comparação entre itens calibrados na prova escrita e no computador.....	70

LISTA DE EQUAÇÕES

Equação 1	- Modelo Logístico de um Parâmetro.....	35
Equação 2	- Modelo Logístico de dois Parâmetros.....	35
Equação 3	- Modelo Logístico de três Parâmetros.....	35
Equação 4	- Função de Informação do Item.....	37
Equação 5	- Função de Informação do Teste.....	37
Equação 6	- Modelo de Resposta Nominal proposto por Bock.....	39
Equação 7	- Modelo de Resposta Gradual proposto por Samejima.....	40
Equação 8	- Modelo de Escala Gradual proposto por Andrich.....	40
Equação 9	- Modelo de Crédito Parcial proposto por Masters.....	40
Equação 10	- Modelo de Crédito Parcial generalizado desenvolvido por Muraki.....	41

LISTA DE ABREVIATURAS E SIGLAS

BPMN	Business Process Modeling Notation
CAT	Testagem Adaptativa Computadorizada
CCI	Curva Característica do Item
CII	Curva de Informação do Item
COLUN	Colégio Universitário
DBA	Database Administrator
EAD	Educação a Distância
EAP	Estimação da Proficiência pela Média da Posteriori
ENADE	Exame Nacional de Desempenho de Estudantes
ENCCEJA	Exame Nacional para Certificação de Competência de Jovens e Adultos
ENEM	Exame Nacional do Ensino Médio
ESL	English as a Second Language
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LSI	Laboratório de Sistemas Inteligentes
MEC	Ministério da Educação
ML1	Modelo Logístico de 1 Parâmetro
ML2	Modelo Logístico de 2 Parâmetros
ML3	Modelo Logístico de 3 Parâmetros
PCN	Parâmetro Curriculares Nacionais
PISA	Programa Internacional de Avaliação de Alunos

SAEB	Sistema de Avaliação da Educação Básica
SARESP	Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo
SAT	Scholastic Assessment Test
SISU	Sistema de Seleção Unificada
SQL	Structured Query Language
TCM	Teoria Clássica de Medidas
TOEFL	Test of English as a Foreign Language
TIC	Tecnologia de Informação e Comunicação
TRI	Teoria de Resposta ao Item
UML	Unified Modeling Language
VT	Virtual-TANEB
Virtual-TANEB	Ambiente Virtual de Avaliação do Sistema de Avaliação da Educação Básica
WEB	World Wide Web

SUMÁRIO

LISTA DE FIGURAS	
LISTA DE TABELAS	
LISTA DE EQUAÇÕES	
LISTA DE ABREVIATURAS E SIGLAS	
1 INTRODUÇÃO	18
1.1 Objetivos do Trabalho	19
1.1.1 Objetivo Geral	19
1.1.1 Objetivos Específicos	19
1.2 Justificativa e Relevância	20
1.3 Estruturação do Trabalho	21
2 APRESENTAÇÃO E DISCUSSÃO DA PROPOSTA	23
2.1 Evolução da Educação	23
2.2 Evolução do Processo de Avaliação	23
2.3 Conclusão	24
3 LEVANTAMENTO BIBLIOGRÁFICO DE TRABALHOS EXISTENTES	25
3.1 Problemas Encontrados	28
3.2 Conclusão	29
4 APLICAÇÃO DA TRI NO PROCESSO DE AVALIAÇÃO E ANÁLISE DE DADOS	30
4.1 Teoria de Medidas para Avaliação Educacional	30
4.1.1 Teoria Clássica de Medidas – TCM	31
4.1.2 Teoria da Resposta ao Item	32
4.2 Uso da TRI na análise de dados	42
4.3 Conclusão	43
5 RECURSOS COMPUTACIONAIS QUE TRABALHAM COM A TRI	44
5.1 Ferramentas	44
5.1.1 PARSCALE	44
5.1.2 SOFTWARE R	44

5.1.3 BILOG – MG.....	45
5.2 Conclusão.....	46
6 MODELAGEM DO SISTEMA.....	47
6.1 Modelagem BPMN do Sistema de Avaliação.....	49
6.2 Diagramas de Caso de Uso do Módulo de Avaliação.....	50
6.3 Conclusão.....	53
7 APLICAÇÃO DA TRI NO PROCESSO DE AVALIAÇÃO.....	54
7.1 Metodologia dos Testes.....	55
7.2 Aplicação do Teste.....	55
7.3 Testes e Resultados do Sistema de Avaliação aplicado.....	59
7.4 Análise dos Resultados.....	71
7.5 Conclusão.....	73
8 CONSIDERAÇÕES FINAIS.....	74
8.1 Contribuições.....	74
8.2 Perspectivas Futuras.....	75
REFERÊNCIAS.....	77
ANEXO.....	79

1 INTRODUÇÃO

O processo educacional o qual nos deparamos hoje exige que todos os alunos possuam uma visão futura, que acompanhe a evolução do processo e dos componentes que compõem o mesmo. O estudante do século XXI deve estar disposto a enfrentar desafios e superar limites.

Assim, com o propósito de avaliar a evolução e o desenvolvimento do aluno, surgiram diversos sistemas de avaliação educacional. Estes sistemas propõem por meio de testes a constante avaliação dos alunos, do ensino oferecido e da infraestrutura escolar. Por meio do resultado, é possível realizar um acompanhamento do ensino, tentando solucionar as falhas detectadas no processo.

No entanto, para que seja feita a elaboração e a avaliação dos testes aplicados, é necessário adotar uma metodologia para avaliação do desempenho dos alunos da forma mais próxima possível da realidade.

As Teorias de Medidas introduzem modelos estatísticos que formam um modo bastante interessante de se analisar a aprendizagem dos alunos e as falhas existentes na estrutura escolar por meio de amostragem (VERAS, 2010).

Assim, visualizando as Teorias de Medidas como uma forma coerente e confiável de avaliação, os processos de avaliação educacional, na sua grande maioria fazem uso desta teoria. A Teoria de Medidas oferece métodos estatísticos de avaliação, se tratando de medidas objetivas (testes) pode basear-se em dois modelos oferecidos pela Teoria de Medidas: A Teoria Clássica de Medidas (TCM) e a Teoria de Resposta ao Item (TRI).

A Teoria Clássica de Medidas (TCM) permite a análise dos itens de uma prova baseando-se em parâmetros descritivos dos itens. Os resultados encontrados dependem do conjunto de itens que compõem o instrumento de medida, ou seja, as análises e interpretações estão sempre associadas à prova como um todo, tornando assim, difícil a comparação entre indivíduos que não foram submetidos às mesmas provas. Na Teoria de Resposta ao Item é possível a comparação entre populações, desde que submetidas a provas que tenham alguns itens comuns, ou ainda, a comparação entre indivíduos da mesma população que tenham sido submetidos a

provas totalmente diferentes, isto é possível devido ao fato de que a TRI tem como elementos centrais os itens e não a prova como um todo (ANDRADE, 2000).

No entanto, o processo de avaliação no qual nos deparamos nos dias atuais, exige um impacto inovador que atribua melhorias ao processo educacional. Neste ponto que a educação se encontra com a inovação tecnológica. O processo educacional necessita de inovação, qualidade, segurança e rapidez, e a automatização desse processo poderá atribuir essas melhorias.

Dessa forma, no decorrer deste trabalho serão apresentadas e utilizadas, metodologias e ferramentas que possibilitam a automatização e agilidade do processo educacional, no entanto, para isso será feita a análise dos diferentes tipos de avaliação a partir do desenvolvimento de um módulo de avaliação web composto pela Teoria de Resposta ao Item, que permita a elaboração e avaliação de testes seguindo os modelos e parâmetros exigidos pela TRI.

Este trabalho de dissertação poderá ser usado como instrumento de experiência em um processo de avaliação em constante mudança. Na análise de dados será utilizada a TRI e a partir dos resultados obtidos, será possível analisar a variação de comportamento entre um processo de avaliação tradicional, aplicado por um caderno de provas no papel e caneta, com a avaliação feita no computador.

1.1 Objetivos do Trabalho

1.1.1 Objetivo Geral

Esta pesquisa tem como objetivo propor um Módulo de Avaliação on-line baseado nos princípios da TRI com a finalidade de aplicá-lo no processo de Transição e Avaliação de provas de caráter diferenciado.

1.1.2 Objetivos Específicos

- Fazer uma revisão bibliográfica de trabalhos existentes e que possuem o mesmo princípio;

- Entender o funcionamento dos processos de avaliação atuantes no Brasil e no mundo;
- Aplicar a TRI no Processo de Avaliação;
- Planejar e modelar um sistema de avaliação que torne possível a realização de testes seguindo os princípios da Teoria de Resposta ao Item;
- Desenvolver o módulo de avaliação e testá-lo;
- Aplicar o teste em duas formas: escrita e computador;
- Justificar os resultados e relatar os problemas encontrados;
- Fazer uma análise comparativa de resultados entre a avaliação tradicional, no papel e caneta, e a avaliação feita no computador.

1.2 Justificativa e Relevância

Os sistemas de avaliação existentes atualmente só são possíveis de serem executados devido à evolução tecnológica e expansão que o ensino teve junto à evolução da sociedade (TEIXEIRA, 2005).

A ideia de empregar o computador no suporte a atividades em educação e treinamento não é recente. Contudo, o panorama mundial atual, globalizado, exige o uso das chamadas Tecnologia de Informação e Comunicação (TIC) e, como não poderia deixar de ser, o sistema educativo não deve ficar de fora desse processo, fazendo-se, entretanto, necessário ao mesmo adaptar-se a esse atual momento (FARIAS, 2010).

Conforme os Parâmetros Curriculares Nacionais (PCN) (PCN,1999), as propostas de mudanças qualitativas para o processo de ensino-aprendizagem indicam a sistematização de um conjunto de disposições e atitudes como pesquisar, selecionar informações, analisar, sintetizar, argumentar, negociar significados, cooperar, de forma que o aluno possa participar do mundo social, incluindo-se aí, a cidadania, o trabalho e a continuidade dos estudos.

Para atingir os objetivos propostos pelo PCN, a Teoria de Resposta ao Item visa sintetizar estes aspectos de forma estatística, buscando resgatar o espírito de pesquisa e aquisição de conhecimento entre os estudantes, já que a aplicação da TRI modela a probabilidade de existir o acerto ao acaso em testes de avaliação.

Nas últimas décadas, a TRI vem tornando-se a técnica predominante no campo de testes em vários países (ANDRADE, 2000). O seu uso no processo de avaliação possibilita que:

- As questões contidas nas provas sejam analisadas individualmente.
- Os graus de dificuldade, discriminação e acerto ao acaso são parâmetros que podem ser colocados nas equações existentes para o cálculo da relação entre a probabilidade de acerto e a habilidade do testando;
- Quando obedecidos os pressupostos exigidos pelo modelo unidimensional, os parâmetros das questões (dificuldade, discriminação e acerto ao acaso) são invariantes, ou seja, são constantes independentes dos respondentes na qual são aplicadas;
- Pessoas com o mesmo número de acerto poderão ter notas diferentes, no entanto o acerto ao acaso independe do grau de dificuldade do item;
- A identificação de acerto ao acaso fica facilitada, já que a probabilidade que um estudante com baixo nível de habilidade acerte questões com alto nível de dificuldade é baixa.

Assim, tem-se que a TRI possui um importante papel no processo de avaliação e por isso optou-se no desenvolvimento deste trabalho, associar os conceitos apresentados pela TRI ao desenvolvimento de um módulo que possibilite a realização de um teste e sua aplicação, como experimento no levantamento de dados que nos possibilite a comparação e a equalização de modelos distintos de provas.

1.3 Estruturação do trabalho

A dissertação aqui desenvolvida encontra-se estruturada em nove capítulos incluindo a introdução.

Os primeiros capítulos abordam pesquisas sobre estudos aplicados já existentes ou em desenvolvimento e que possuem objetivos semelhantes aos apresentados nesta dissertação, assim tem-se que:

No capítulo 2, discute-se sobre a importância da proposta apresentada com base em fundamentações pedagógicas que justificam a inovação tecnológica no processo educacional.

No capítulo 3, faz-se uma análise bibliográfica nos trabalhos existentes que exploram objetivos iguais ou semelhantes ao objetivo proposto e são apresentados os problemas encontrados.

No capítulo 4, com base no estudo realizado, apresenta-se de que forma as Teorias de Medidas são usadas no processo de avaliação educacional e como a TRI poderá ser usada para avaliar a comparação e o impacto causado entre os dois tipos de prova (escrita e no computador).

No capítulo 5, apresenta-se a pesquisa feita com os recursos computacionais que possuem parâmetros possíveis de se trabalhar na TRI, assim foi utilizado um recurso computacional na análise dos dados apresentados nesta dissertação.

No capítulo 6, explicita-se os experimentos propostos, assim neste capítulo foi feita a modelagem do banco e a modelagem BPMN do processo de avaliação, o que tornou possível o desenvolvimento do módulo de avaliação aqui proposto.

No capítulo 7, apresenta-se de que forma a prova foi aplicada e como a TRI foi utilizada na análise dos dados, apresentando assim, os resultados encontrados.

E finalmente, no capítulo 8, é apresentada a conclusão, os resultados alcançados, as principais contribuições e pesquisas futuras.

2 APRESENTAÇÃO E DISCUSSÃO DA PROPOSTA

A proposta apresentada reúne conceitos referentes à educação, processo de avaliação educacional, teoria de medidas, ferramentas estatísticas dentre outros. Estes conceitos possibilitam que vários temas sejam explorados, é o caso do trabalho aqui apresentado.

Este trabalho foi iniciado a partir da ideia de desenvolver um módulo de avaliação que utilizasse a TRI na análise de dados. À medida que este módulo foi sendo desenvolvido surgiram novas ideias, e uma delas é que o uso deste módulo desenvolvido associado à TRI pudesse ser utilizado para explorar o processo de avaliação educacional.

Foi a partir de então que se resolveu utilizar o módulo de avaliação desenvolvido e a TRI na análise do processo de migração da prova escrita, feita no papel e lápis, para a prova no computador.

2.1 Evolução da Educação

De acordo com Teixeira (2005), a evolução tecnológica ocorrida nos últimos tempos ressaltou a necessidade de introduzir novas técnicas no ensino e o uso de computador na educação mostrou-se muito interessante. A Internet possibilitou uma nova dimensão ao processo ensino-aprendizagem, e o computador aliado à evolução das telecomunicações, da telemática e da Internet, redimensiona o processo educativo, ele rompe as barreiras do espaço e do tempo, aumentando as possibilidades educativas e o avanço da Educação a distância (EAD), hoje, inevitavelmente mediada pelas Tecnologias de Informação e Comunicação (TIC), que ganharam resignificação, a partir do uso do computador e da Internet.

2.2 Evolução do Processo de Avaliação

Atualmente a avaliação carrega consigo todos os conceitos até então aplicados, atribuindo novos paradigmas e reformulando valores a fim de adequá-los à época; ou seja, o processo de avaliação continua obtendo o mesmo princípio, o que muda são as inovações e aplicabilidade do processo (FARIAS, 2011).

Assim, o conhecimento continua sendo a base para a formulação de um processo de avaliação educacional, no entanto ressalta-se que, além do conhecimento como base, a integração das novas tecnologias com os meios educacionais, torna-se cada vez mais indispensável ao processo (MACHADO, 1996).

O processo de avaliação educacional está em constante busca por alternativas que melhorem o diagnóstico escolar, não apenas em relação ao conhecimento do estudante e do professor mais também à infraestrutura oferecida pela instituição, assim começou o despertar para as Teorias de Medidas (TCM e TRI), visto que elas forneceria uma medida na busca pelo diagnóstico escolar (VERAS, 2010).

Baseando-se nesses conceitos e informações, despertou-se a necessidade de além de desenvolver um módulo de avaliação e aplicar a TRI na análise de dados, verificar o impacto que a mesma prova aplicada de forma escrita e no computador poderá causar processo de avaliação aplicado.

2.3 Conclusão

No decorrer deste capítulo, foram apresentadas por meios de fundamentações teórico-pedagógicas, a necessidade da evolução do processo de avaliação, visto que a educação está em processo de transformação e ao longo dos próximos anos será inevitável integrar a tecnologia a este processo. Justificando assim o objetivo central da proposta, que é a inserção da tecnologia no processo de avaliação e a análise do impacto causado pela migração do papel e lápis para o computador.

3 LEVANTAMENTO BIBLIOGRÁFICO DE TRABALHOS EXISTENTES

A educação ao longo dos anos vem ganhando destaque no que se refere à evolução, principalmente nas formas adquiridas para aquisição do conhecimento. Há muito tempo a avaliação é usada como um instrumento de medida, ela serve para medir o grau de conhecimento adquirido pelo estudante em determinada área de ensino. No entanto, as formas de avaliar estão numa escala crescente, e com a inserção das novas tecnologias no âmbito educacional, uma nova teoria de medida, a TRI, surgiu e trouxe consigo uma ampla bagagem de diversas possibilidades de avaliação, em especial ao sistema de testagem adaptativa computadorizada (CAT).

Neste trabalho não se utilizou a Testagem Adaptativa Computadorizada, no entanto o computador foi utilizado como instrumento na coleta de dados para análise via TRI.

No mundo em que a tecnologia cresce gradativamente, sentiu-se a necessidade de transformar hábitos e costumes em algo mais próximo da realidade, que é a inserção gradativa da tecnologia no âmbito educacional, por isso a ideia de migrar a avaliação tradicionalmente conhecida no papel e lápis, para um modelo de avaliação que se encaixe no crescimento tecnológico educacional, a avaliação feita no computador.

Existem alguns trabalhos realizados ou em desenvolvimento, que assim como o trabalho aqui desenvolvido, utilizam um módulo educacional computadorizado na coleta de dados associado a TRI na análise dos mesmos, existem ainda trabalhos que avançam um pouco mais na tecnologia, e adotam a CAT como modelo de avaliação.

É o que se pode constatar nas citações a seguir:

- As Novas Tecnologias e a Educação:

[...] O desenvolvimento técnico-científico, por sua vez, impulsionando novas descobertas, gera grandes alterações na vida humana e no trabalho, caracterizando este momento como período da Terceira Revolução Industrial ou Revolução Tecnológica. A exigência de um domínio cada vez

maior de conhecimentos e habilidades, para tratar desta realidade diversa e complexa, impõe novas concepções de educação, escola e ensino.

As tecnologias da informação ou novas tecnologias da informação e comunicação são o resultado da fusão de três vertentes técnicas: a informática, as telecomunicações e as mídias eletrônicas. Elas criaram um encantamento no meio educacional...

[...] A Web tem sido uma fonte acessível para pesquisa; através dela se tem acesso a bibliotecas, banco de dados, fóruns, possibilidade de conversações telefônicas e de videoconferências. (PINTO, 2004).

De acordo com Pinto (2004), ele trata da importância da tecnologia no Processo Educacional, e abre caminho para novos experimentos relacionados à evolução desse processo. A inserção das tecnologias trata de mudanças diretamente realizadas no meio educacional, no quais os alunos buscam apoio para aquisição do conhecimento, nas ferramentas oferecidas pela tecnologia, assim a migração da avaliação escrita para a avaliação no computador começa a se tornar viável aos meios educacionais de avaliação.

- Análise de Itens de uma prova de raciocínio estatístico:

A TRI não entra em contradição com os princípios da psicometria clássica, e traz uma nova proposta estatística, a de análise centrada nos itens, que supera as limitações da teoria clássica, além de apresentar novos recursos tecnológicos para a avaliação psicológica e educacional (Primi, 1998). Existem vários modelos possíveis de resposta ao item, que diferem em sua forma em função da característica do item e do número de parâmetros especificados no modelo. Todos os modelos dessa teoria possuem um ou mais parâmetros que descrevem o item e um ou mais parâmetros que descrevem o respondente. O primeiro passo para uma aplicação da TRI é a estimação desses parâmetros (HAMBLETON, 1991).

Alguns pontos tem sido levantados na literatura sobre a adequação desta teoria na área de avaliação educacional. Dois deles considerados importantes são: a dimensionalidade do espaço de traços latentes envolvidos na avaliação e a equalização de diferentes avaliações. É necessário ressaltar que, apesar de não haver dúvidas de que a aplicação desta teoria muito contribui para a melhoria das avaliações educacionais em geral, sua disseminação efetiva depende da integração de especialistas das áreas de estatística e educação (Andrade, 2001).

Uma das vantagens da utilização da TRI na avaliação educacional é que esta possibilita análises qualitativas a partir dos resultados brutos de uma prova, fornecendo assim informações mais precisas do desempenho dos respondentes e da qualidade das questões utilizadas (itens), questões que devem ter índices de dificuldade e de discriminação aceitáveis e correlacionadas com a prova total. (VENDRAMINI, 2004).

Os referidos autores fazem um experimento em que uma prova de estatística é aplicada, fazendo uso da Teoria Clássica de Medidas e da Teoria de Resposta ao Item. Dessa forma, com os resultados coletados, fazem um

levantamento das vantagens e desvantagens de se usar a TRI no processo de avaliação.

- Sistema de Testagem Adaptativa Computadorizada (CAT) da Universidade de Brasília com Pré-Testagem Automatizada:

A Universidade de Brasília (UnB) tem realizado, desde o segundo semestre de 2010, exames de proficiência em Língua Inglesa administrados por computadores, utilizando uma moderna tecnologia, ainda pouco empregada em instituições brasileiras, chamada de Teste Adaptativo Computadorizado ou, em inglês, Computerized Adaptive Test (CAT). Tal tecnologia consiste em um teste computadorizado que se adapta segundo o nível de proficiência do avaliando em tempo real, de forma a estimar seu nível com maior precisão do que testes tradicionais.

[...] Consta ainda neste trabalho um estudo sobre a teoria de avaliação utilizada neste sistema, chamada Teoria de Resposta ao Item (TRI), uma discussão sobre os processos utilizados na aplicação de tal avaliação, e uma avaliação sobre os estados do sistema antes e depois da adição das novas funcionalidades.

[...] Deve se fazer claro que um CAT não necessariamente precisa usar, como método de avaliação, a TRI. Contudo, este foi o método usado pelo CESPE/UnB para a construção do CAT.

[...] Uma das exigências da TRI é que os itens que compõem este banco estejam calibrados. Para calibrar um item, este precisa passar pelo processo chamado de pré-testagem. A pré-testagem é a atividade de coletar uma quantidade pré-determinada de respostas a um item, de um ou mais grupos de avaliandos. A partir desta coleta, a TRI é capaz de estimar os parâmetros psicométricos do item, ou seja, calibrá-lo.

[...] Do ponto de vista da avaliação, um teste adaptativo, a princípio, não tem como necessidade o uso de computadores. As questões poderiam ser selecionadas e apresentadas ao avaliando por um aplicador humano. Contudo, para um grande número de avaliandos, essa opção obviamente se tornaria cada vez mais inviável. O uso de computadores resolve esse problema (CARVALHO, 2011).

No trabalho dos autores supracitado, o fator esclarecedor em questão é o fato de poder se trabalhar com o CAT e a TRI juntas, além de levar em consideração a vantagem de fazer um teste no computador, já que este tipo de avaliação não exige o uso do computador, mais favorece e torna-se necessária ao processo de avaliação, quando se trata de grande número de avaliandos.

- O Efeito dos Computadores no Teste e Confiabilidade de Inter-Avaliadores de Testes Escritos de Aprendizes de ESL

[...] Desde os anos 1970, os computadores tem tido na escola e em casa uma influência considerável sobre educação (Zandvliet e Farragher, 1997). Assim, por três décadas, os teóricos da educação e pesquisadores tem proposto várias maneiras em que o computador influencia a educação. Como resultado dessa

influencia, nos últimos anos tem existido uma explosão de interesse quanto ao uso de computadores.

O uso do computador em testes de linguagem ganhou destaque quando tornou possível por meio da TRI, os testes individuais. Os avanços na TRI juntamente com a informática desempenharam importante papel no desenvolvimento de testes de linguagem na década de 1990. Assim, os computadores passaram a ser aceitos em métodos avaliativos.

[...] Os pesquisadores argumentaram que o uso do computador ajuda os alunos a prevenir a ansiedade sobre a escrita. No entanto, poucas pesquisas sobre o uso do computador apareceram com escritores da segunda língua, apesar de muitos estudos sobre a escrita ter sido realizada por falantes nativos. Poucos estudos sobre a linguagem escrita, mostraram que escritores da segunda língua são frequentemente assumidos a ter mais apreensão do que os escritores de língua nativa. Por outro lado, de acordo com alguns estudos, a utilização do computador parece ter tido efeitos positivos sobre escritores da segunda língua (Phinney, 1991), embora o nível de pesquisa da linguagem escrita e no computador não chega perto da atividade realizada por escrito quando se refere a primeira língua.

[...] Por exemplo, de acordo com Phinney e Mathis (1991), os aprendizes de segunda língua sentiram que o computador melhorou a atitude em relação à escrita em Inglês. Os alunos também pareciam passar mais tempo escrevendo que os alunos que não usaram um computador, além de terem produzido trabalhos mais longos (AYDIN, 2006).

Os resultados porém apresentados neste trabalho visou avaliar os testes de Inglês feitos na modalidade escrita e no computador, fazendo uso ainda da TRI no processo de avaliação.

3.1 Problemas Encontrados

Conforme Pinto (2004), para a realização da Evolução da Tecnologia no processo educacional, é necessário discutir a chamada sociedade da informação, como ela pode ser articulada no repensar da educação pública, como os elementos da sociedade da informação e comunicação estão influenciando o processo educativo e realizar uma reflexão sobre o comportamento da escola pública diante da proposta de “Informatização”.

Assim, de acordo com análise e reflexão proposta pelo autor acima, nota-se que a necessidade da inserção de novas tecnologias ao processo educacional existe, no entanto a educação oferecida nas escolas públicas deve ser repensada a fim de oferecer aos alunos um contato maior com a tecnologia na escola.

Este fato, em que os alunos não possuem contato com a tecnologia na escola é algo possível de ser avaliado no decorrer deste trabalho de dissertação.

Na inserção da TRI no processo de avaliação educacional, os resultados encontrados, apesar de terem sido satisfatórios mostrando que quando comparados

os índices de facilidade de TCM com os de dificuldade da TRI, observou-se quanto maior a proporção de acertos no item, é menor o seu índice de dificuldade. Estes resultados indicaram a necessidade de outros estudos com a inclusão de novos itens na prova, para se chegar a uma prova que apresente índices mais elevados de confiabilidade e validade a ser utilizada no contexto educacional.

E por fim, o resultado do impacto causado pela migração de diferentes modalidades de teste, como afirma Aydin (2006), foi conflitante e não conclusivo. Já que, alguns estudos mostraram que ocorreram mais notas baixas na versão feita no computador, e outros estudos mostraram que não houve diferença entre as versões digitais e manuscritas no processo de classificação.(AYDIN, 2006).

3.2 Conclusão

No decorrer deste capítulo, conforme os trabalhos apresentados, observou-se a importância que a tecnologia tem na educação, e o quão interessante torna-se a parceria entre a TRI e o computador.

4 APLICAÇÃO DA TRI NO PROCESSO DE AVALIAÇÃO E ANÁLISE DE DADOS

A TRI vem sendo constantemente aplicada em processos de avaliação. Neste trabalho, os testes foram aplicados de duas formas: escrita manual e no Computador, e a TRI foi aplicada na análise dos resultados dos dois modelos de teste apresentados. Ela possui diversos conceitos e modelos que se aplicam neste tipo de avaliação e para uma melhor compreensão da Teoria de Resposta ao Item e sua aplicação neste tipo de análise serão apresentadas as Teorias de Medidas existentes no processo de avaliação educacional.

4.1 Teoria de Medidas para Avaliação Educacional

As Teorias de Medidas são utilizadas em processos de avaliação para obtenção de resultados em diversos testes, e são muito utilizadas na representação de desempenho de alunos.

Segundo Andrade, Tavares e Valle (2000), as Teorias de Medidas possuem diversas características que demonstram de que forma elas se aplicam nos processos de avaliação. Existem características que são comuns às duas teorias, a TCM e a TRI. Observe que, na Teoria Clássica de Medidas, os resultados extraídos de uma prova são expressos por seus escores brutos ou padronizados e por isso são comumente utilizados nos processos de avaliação e seleção de indivíduos, porém estes resultados, diferentemente de como acontece na TRI, dependem de conjunto de itens, que compõem o instrumento de medida, ou seja, as análises e interpretações estão sempre associadas à prova como um todo, enquanto que, na TRI as análises e interpretações se concentram nos itens e não no conjunto de itens.

Dessa forma, os mesmos autores relatam que diferente de como ocorre na TCM e a partir das características apresentadas pela TRI, é possível que na aplicação prática de um teste, sejam avaliados o desenvolvimento de diferentes grupos de alunos que se submeteram a diferentes testes, ou seja, a comparação

entre populações submetidas a provas com alguns itens comuns, ou a comparação entre indivíduos da mesma população que tenham sido submetidos a provas totalmente diferentes

4.1.1 Teoria Clássica de Medidas - TCM

A Teoria Clássica de Medidas (TCM) possui seus escores brutos ou padronizados e os seus resultados dependem do conjunto de itens que compõem o instrumento de medida, e por isso é inviável que através dessa teoria possa-se fazer análises na comparação de resultados entre indivíduos que não foram submetidos ao mesmo instrumento de medidas (ANDRADE, 2000).

Segundo Tezza e Bornia (2009), a TCM dispõe de uma equação básica que descreve a relação entre os escores observados, os escores verdadeiros e o erro. Observe a equação abaixo:

$$X = T + E$$

em que: X é o escore observado;

T, o escore verdadeiro;

E, o erro de medida.

O escore verdadeiro (T) representa o traço latente e o erro (E) pode estar associado a alguns fatores:

- Impossibilidade de incluir um número infinito de itens no conjunto de itens;
- Impossibilidade de aplicar o conjunto de itens um número infinito de vezes;
- Impossibilidade de submeter o conjunto de itens a um número infinito de organizações.

Conforme Reise (2005 *apud* TEZZA, 2009). Na Teoria Clássica de Medidas, as informações do item, a confiabilidade dos resultados e os erros-padrões, são dependentes das características particulares da população. Dessa forma, o (X) em questão, que é o escore observado é definido por um conjunto particular de itens em uma simples medida.

A elaboração de um instrumento de medida com o uso da Teoria Clássica de Medidas tem os parâmetros dos itens dependentes das amostras os quais eles foram calculados. O parâmetro de dificuldade do item vai variar conforme a pesquisa

realizada em função da amostra, assim, de acordo com a amostra, o item se tornará mais fácil ou mais difícil. Na TCM, a estimação do parâmetro se baseia no escore total de um teste, isso supõe que todos os itens sejam adequados, mais isto é uma inverdade. Conforme algumas situações já observadas, a suposição de que a variação dos erros de medida (E) é igual para todos os respondentes é algo difícil de ser sustentado (HAMBLETON 1985 *apud* TEZZA, 2009).

Tendo em vista estas limitações, Pasquali (2003 *apud* TEZZA, 2009), argumentam que, dentro da Teoria Clássica de Medidas, é difícil realizar comparações entre grupos e organizações, visto que, não é possível obter da mesma forma o escore total de um conjunto de itens fáceis e o escore total de um conjunto de itens difíceis. Para tentar solucionar estes e outros problemas encontrados na TCM tem sido realizado um trabalho mais profundo com a Teoria da Resposta ao Item e tem apresentado várias vantagens em relação à TCM.

Uma das grandes vantagens da TRI sobre a Teoria Clássica é que ela permite a comparação entre populações, desde que submetidas a provas que tenham alguns itens comuns, ou ainda, a comparação entre indivíduos da mesma população que tenham sido submetidos a provas totalmente diferentes. Isto porque uma das principais características da TRI é que ela tem como elementos centrais os itens e não a prova como um todo (ANDRADE, 2000).

4.1.2 Teoria da Resposta ao Item – TRI

De acordo com Andrade, (2000) no âmbito da avaliação educacional é comum que se defina a população conforme características que podem variar ou não, dependendo dos objetivos do estudo podendo ser relevantes ou não para diferenciação de populações.

Ressaltam os autores que a Teoria da Resposta ao Item (TRI) se define por um conjunto de modelos matemáticos e estatísticos que podem ser utilizados na análise de itens e escalas, para criar e administrar medidas e medir indivíduos ou organizações. Ela é sustentada basicamente por três fundamentos: Função resposta de um item, função de informação e invariância. Ressaltam ainda os autores que, a TRI é definida como um conjunto de modelos matemáticos que procuram

representar a probabilidade de um indivíduo apontar uma certa resposta a um item como função dos parâmetros do item e da habilidade do respondente. Essa relação é expressa de forma que, quanto maior a habilidade, maior a probabilidade de acerto no item. Existem modelos cumulativos e modelos de desdobramentos. Os modelos propostos na literatura dependem de três fatores principais:

- Da natureza do item – dicotômico ou não dicotômico (um item binário, ou seja admite duas respostas possíveis - certo e errado);
- Do número de populações envolvidas – apenas uma ou mais de uma;
- Da quantidade de traços latentes que está sendo medida – apenas um ou mais de um.

No entanto, Vendramini (2004) afirma que a escolha do modelo mais adequado depende daquilo que se deseja medir e da natureza dos dados disponíveis. A literatura dispõe de modelos unidimensionais e multidimensionais, neste caso será utilizado o Modelo Unidimensional e para isso, é necessário seguir alguns critérios para sua utilização:

- Critério da unidimensionalidade: os itens de um teste devem medir uma única habilidade, ou ao menos deve haver um fator dominante que influencie o desempenho dos respondentes no teste;
- Critério da independência local: as respostas dos participantes aos itens não devem ser influenciadas pelas respostas fornecidas a outros itens, para indivíduos com uma mesma habilidade.

Os parâmetros dos itens mais relevantes são: a dificuldade, a discriminação e a probabilidade de acerto por acaso. A dificuldade do item é dada na mesma escala da habilidade, ou seja, à habilidade necessária que um indivíduo deve ter para uma dada probabilidade de acertar o item. A discriminação indica a variação correspondente à probabilidade de indivíduos de diferentes habilidades acertarem um item, e o acerto ao acaso é a probabilidade de um sujeito com baixa habilidade acertar um item.

A inclinação da Curva Característica do Item – CCI corresponde à discriminação. O eixo das abscissas indica o traço latente do indivíduo e é representado por θ , cujo valor pode variar de $-\infty$ a $+\infty$, e no eixo das ordenadas a probabilidade de responder corretamente ao item, podendo variar de 0 a 1.

Na TRI, a probabilidade de acertar um item depende da habilidade do indivíduo, que permite expressar a sua resposta aos itens, e dos parâmetros dos itens (a_i , b_i e c_i), por este motivo é necessário estimar valores destes parâmetros que expliquem os resultados obtidos, com base nas respostas dos participantes aos itens. Esse processo é chamado de calibração ou parametrização, e é feito com o auxílio de programas específicos como: RASCAL, XCALIBRE, BILOG, Software R, entre outros (VENDRAMINI, 2004).

Modelos Matemáticos

A TRI dispõe de vários modelos matemáticos que foram desenvolvidos para resolver problemas inerentes às diversas situações e itens de diferentes naturezas. A tabela 1 apresenta os modelos que serão estudados no decorrer do trabalho:

Tabela 1 - Modelos Matemáticos

Natureza do Item	Modelo
Dicotômico	Modelo Logístico de 1 Parâmetro ou Modelo de Rasch (<i>One-Parameter Logistic Model or Rasch Model</i>) Modelo Logístico de 2 Parâmetros (<i>Two-Parameter Logistic Model</i>) Modelo Logístico de 3 Parâmetros (<i>Three-Parameter Logistic Model</i>) Modelo Logístico de 4 Parâmetros (<i>Four-Parameter Logistic Model</i>)
Politômico ou Não dicotômico	Modelo de Resposta Nominal (<i>Nominal Categories Model</i>) Modelo de Resposta Gradual (<i>Graded Response Model</i>) Modelo de escala Gradual (<i>Rating Scale Model</i>) Modelo de Crédito Parcial (<i>Partial Credit Model</i>) Modelo de Crédito Parcial Generalizado (<i>Generalized Partial Credit Model</i>)

- Modelo Logístico de um Parâmetro (ML1):

O modelo de um parâmetro, ou Modelo de *Rasch* ilustrada na equação, avalia em sua expressão o índice de dificuldade, representado por b_i , e dispõe de variáveis que verificam a probabilidade de um indivíduo acertar o item i de forma casual. (VENDRAMINI, 2004).

$$P(X_{ij} = 1 | \theta_j) = P_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}}$$

Equação 1 - Modelo Logístico de um Parâmetro (ANDRADE, 2000)

- Modelo Logístico de dois Parâmetros (ML2):

O Modelo de dois Parâmetros avalia em sua equação o índice de dificuldade, representado por b_i e de discriminação, representado por a_i , dispõe de variáveis que influenciam na probabilidade de um indivíduo acertar o item i ao acaso. (VENDRAMINI, 2004).

$$P_i(\theta_j) = P(U_{ij} = 1 | \theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}}$$

Equação 2 - Modelo Logístico de dois parâmetros (ANDRADE, 2000)

- Modelo Logístico de três Parâmetros (ML3):

O modelo de três parâmetros é o modelo mais completo, visto que, ele avalia em sua expressão além dos parâmetros vistos nos modelos anteriores, que são: a_i (parâmetro de discriminação) e b_i (parâmetro de dificuldade), a probabilidade de acerto ao acaso que é representada por c_i . Este modelo é representado pela seguinte equação matemática (VENDRAMINI, 2004).

$$P(X_{ij} = 1 | \theta_j) = P_i(\theta) = c_i + \frac{(1 - c_i)}{1 + e^{-Da_i(\theta - b_i)}} \quad (1)$$

com $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, m$, sendo:

Equação 3 - Modelo Logístico de três Parâmetros (ANDRADE, 2000)

- X_{ij} : Uma variável dicotômica que assume os valores 1 ou 0, conforme o indivíduo j responda correta ou incorretamente o item i , respectivamente;
- θ_j : O valor que representa a variável latente (aptidão ou habilidade) que permite explicar a resposta do j -ésimo indivíduo aos itens;
- $P(X_{ij} = 1 \mid \theta_j) = P_i(\theta_j)$: É a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i , é chamada de Função Resposta do Item (FRI);
- c_i : A probabilidade de acerto ao acaso;
- b_i : O índice de dificuldade (ou parâmetro de posição) do item i , medido na mesma escala da habilidade θ . Corresponde à habilidade necessária para uma probabilidade de acerto igual a $(1 + c_i) / 2$;
- a_i : O índice de discriminação (ou parâmetro de inclinação) do item i , com valor proporcional à inclinação da CCI no ponto b_i ;
- D : Um fator de escala constante, igual a 1 ou a 1,7 (quando se deseja que a função logística se aproxime da ogiva normal);
- e : Um número transcendental, base dos logaritmos neperianos, cujo valor é aproximadamente 2,718;
- n : O número de itens;
- m : O número de indivíduos.

Curva Característica do Item – CCI

A Curva Característica do Item apresenta em sua análise o índice de proficiência individual dos respondentes conforme apresentado na figura 1.

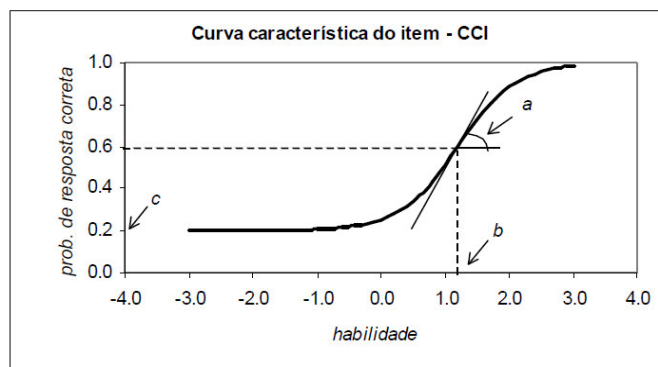


Figura 1 - Curva Característica do Item (ANDRADE, 2000)

O parâmetro **b** representa a habilidade necessária para uma probabilidade de acerto igual a $(1 + c) / 2$. Assim, quanto maior o valor de *b*, maior terá que ser a proficiência exigida para o respondente acertar o item.

O parâmetro **c** representa o acerto ao acaso, ou seja, a probabilidade de um aluno com baixa habilidade responder corretamente o item (ANDRADE, 2000).

Conforme Andrade (2000), o parâmetro **a** define a discriminação do item. Neste modelo itens com **a** negativo não são esperados, visto que, isto indicaria que a probabilidade de responder corretamente o item diminui conforme o aumento da habilidade. Quando o **a** possui valores baixos isto significa que, o item avaliado tem pouco poder de discriminação, assim, alunos com habilidades bastante diferentes tem a mesma probabilidade de responder corretamente ao item. Caso contrário, quando o **a** possui valores altos, isto significa que os alunos podem ser divididos em dois grupos: Com habilidades abaixo do valor do parâmetro **b** e outro grupo cuja habilidade está representada acima do valor do parâmetro **b**. Neste caso a CCI é bastante íngreme.

Função de Informação do Item

Permite analisar quanto um item contém de informação para a medida de habilidade.

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)}$$

Equação 4 - Função de Informação do Item (ANDRADE, 2000)

Função de Informação do Teste

Soma das Informações fornecidas por cada item que compõe o teste.

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

Equação 5 - Função de Informação do Teste (ANDRADE, 2000)

Equalização

É o procedimento para ajustar a medida de habilidade (θ) entre grupos de indivíduos submetidos a diferentes testes com itens em comum para uma mesma métrica (KOLEN *apud* ANJOS, 2012).

De acordo com Andrade (2000), equalizar significa equiparar, tornar comparável, o que no caso da TRI significa colocar parâmetros de itens vindos de provas distintas ou habilidades de respondentes de diferentes grupos, na mesma métrica, isto é, numa escala comum, tornando os itens e/ou as habilidades comparáveis.

Os referidos autores acrescentam que existem dois tipos de Equalização: a equalização via população, em que um único grupo de respondentes é submetido a provas distintas, e a equalização via itens comuns. Na equalização via população, todos os itens devem ser calibrados conjuntamente para ter a garantia de que todos estarão na mesma métrica, já na equalização via itens comuns, a garantia de que as populações envolvidas terão seus parâmetros em uma única escala será dada pelos itens comuns entre as populações.

Os mesmos autores apontam que existem diversas situações as quais a equalização se aplica:

- Equalização via população - um único grupo fazendo duas provas totalmente distintas; um único grupo fazendo duas provas parcialmente distintas;
- Equalização via itens comuns - dois grupos fazendo uma única prova; dois grupos fazendo duas provas parcialmente distintas.

Confirmam ainda os autores que para fazer a equalização é necessário que os itens sejam calibrados. No entanto, a equalização entre duas ou mais populações pode ser feita durante o próprio processo de estimação dos parâmetros.

Também é possível fazer a equalização a posteriori, ou seja, depois de terminado o processo de calibração dos itens.

Segundo Andrade, (2000) a equalização a posteriori calibra-se separadamente os dois conjuntos de itens, que foram submetidos às duas populações de interesse. A condição necessária é que haja itens comuns entre os dois conjuntos. Para itens comuns, há dois conjuntos de estimativas, cada uma na métrica de suas respectivas populações. Daí, através dessas duas estimativas para os itens comuns estabelece-se algum tipo de relação que permita colocarmos os parâmetros de um dos conjuntos de itens na mesma escala do outro. Com todos os itens na mesma métrica, pode-se então estimar as habilidades de todos os respondentes, que então estarão também na mesma escala.

- **Modelo de Resposta Nominal**

Modelo proposto por Bock (1972 *apud* ANDRADE, 2000). Segundo os autores, o propósito deste modelo é maximizar a precisão da habilidade estimada usando toda a informação contida nas respostas dos indivíduos, e não apenas se o item foi respondido corretamente ou não. Este modelo assume que não há nenhuma ordenação, e é representado por:

$$P_{i,k}(\theta_j) = \frac{e^{a_{i,k}^+(\theta_j - b_{i,k}^+)}}{\sum_{h=1}^{m_i} e^{a_{i,h}^+(\theta_j - b_{i,h}^+)}}$$

Equação 6 - Modelo de Resposta Nominal proposto por Bock (ANDRADE, 2000).

- **Modelo de Resposta Gradual**

Modelo proposto por Samejima (1969 *apud* ANDRADE, 2000). Conforme estes autores, assim como no Modelo de Resposta Nominal, este modelo tenta obter mais informação das respostas dos indivíduos do que simplesmente se eles deram respostas corretas ou incorretas.

Os autores afirmam que este modelo pode ser aplicado para itens dicotômicos ou não dicotômicos. No caso de modelos aplicados para itens dicotômicos o parâmetro de inclinação do item pode ser chamado de discriminação

do item. No caso de modelos aplicados para itens não dicotômicos, a discriminação de uma categoria específica de resposta depende tanto do parâmetro de inclinação comum a todas as características do item, quanto da distancia das categorias de dificuldade adjacentes. O referido modelo assume que as características de resposta de um item podem ser ordenadas entre si, representado por:

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k+1})}}$$

Equação 7 - Modelo de Resposta Gradual proposto por Samejima (ANDRADE, 2000).

- **Modelo de Escala Gradual**

Modelo proposto por Andrich (1978 *apud* ANDRADE, 2000). O Modelo de Resposta Gradual definido por Samejima (visto no item anterior) é o modelo de escala gradual. No entanto, aqui é feita uma suposição a mais: a de que os escores das categorias são igualmente espaçados. Este modelo segundo autores acima, também é adequado para itens com categoria de resposta ordenada, e é representado por:

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i + d_k)}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_i + d_{k+1})}}$$

Equação 8 - Modelo de Escala Gradual proposto por Andrich (ANDRADE, 2000).

- **Modelo de Crédito Parcial**

Modelo desenvolvido por Masters (1982 *apud* ANDRADE, 2000). O Modelo de Crédito Parcial possui o mesmo propósito que os demais, no entanto, ele difere do gradual por pertencer à família de modelos de Rasch. Ele é uma extensão do modelo de Rasch para itens dicotômicos. Logo, todos os parâmetros no modelo são de locação, sendo que o poder de discriminação é assumido ser comum para todos os itens.

Este modelo segundo os autores é também um modelo para análise de respostas obtidas de duas ou mais categorias ordenadas.

$$P_{i,k}(\theta_j) = \frac{\exp \left[\sum_{u=0}^k (\theta_j - b_{i,u}) \right]}{\sum_{u=0}^{m_i} \exp \left[\sum_{v=0}^u (\theta_j - b_{i,v}) \right]}$$

Equação 9 - Modelo de Crédito Parcial proposto por Masters (ANDRADE, 2000).

- **Modelo de Crédito Parcial Generalizado**

Modelo desenvolvido por Muraki (1992 *apud* ANDRADE, 2000). Este modelo baseou-se no Modelo de Crédito Parcial de Masters, relaxando a hipótese de poder de discriminação uniforme para todos os itens. Tal modelo é dado por:

$$P_{i,k}(\theta_j) = \frac{\exp \left[\sum_{u=0}^k Da_i(\theta_j - b_{i,u}) \right]}{\sum_{u=0}^{m_i} \exp \left[\sum_{v=0}^u Da_i(\theta_j - b_{i,v}) \right]}$$

Equação 10 - Modelo de Crédito Parcial Generalizado proposto por Muraki (ANDRADE, 2000).

Estimação dos Parâmetros dos Itens e Habilidades

Segundo Andrade (2000), uma das etapas mais importantes da TRI é a estimação dos parâmetros dos itens e das habilidades dos respondentes. Como foi visto no capítulo anterior, a probabilidade de uma resposta correta a um determinado item depende somente da habilidade do indivíduo e dos parâmetros que caracterizam o item. Mas, em geral, ambos são desconhecidos. Apenas as respostas dos indivíduos aos itens do teste são conhecidas.

Assim, 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;
- Quando conhecemos as habilidades dos respondentes, estaremos interessados apenas na estimação dos parâmetros dos itens;
- Quando necessita-se 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. Geralmente as estimações dos parâmetros dos itens e/ou proficiências são feitas pelo Método da Máxima Verossimilhança e também utilizando-se aplicação de algum processo iterativo, como o algoritmo Newton-Raphson e, ainda, o “Scoring” de Fischer. Alguns procedimentos bayesianos também são aplicados com bastante frequência. As soluções que exigem procedimentos iterativos envolvem cálculos bastante complexos e, conseqüentemente, programas computacionais específicos. Em qualquer um desses casos, os valores das proficiências (θ) e dos parâmetros dos itens (a_i , b_i , c_i) estão todos na mesma escala de medida (FRANCISCO, 2005).

Estimação Bayesiana

Consiste em estabelecer distribuições a priori para os parâmetros de interesse, construir uma nova função denominada distribuição a posteriori e estimar os parâmetros de interesse com base em alguma característica dessa distribuição.

4.2 Uso da TRI na análise de Dados

A TRI – Teoria da Resposta ao Item é um método de avaliação que se propõe a avaliar o nível de conhecimento do candidato (proficiência). Ela avalia os itens da prova, diferentemente das provas elaboradas a partir da teoria clássica dos [testes](#), que avaliam a prova como um todo focando na quantidade de acertos (BRASIL, 2012).

Conforme a TRI, um item é considerado bom quando ele diferencia na sua análise os alunos com desempenho superior, daqueles com desempenho inferior, para isso é apresentado um valor positivo e alto do coeficiente de correlação bisserial, que mede a associação quando comportamento empírico de um bom item

é tal que a resposta dada pelos alunos a este item está associada positivamente ao escore total dos alunos no teste. O escore total dos alunos é definido pelo número de itens que o aluno acerta (SANTOS, 2012).

Durante um processo de análise de dados, a TRI possibilita análises qualitativas a partir de resultados brutos de uma prova, fornecendo assim informações mais precisas do desempenho dos respondentes e da qualidade das questões utilizadas, ou seja, os itens (VENDRAMINI, 2004).

Assim, por meios de métodos estatísticos e modelos matemáticos, a TRI alcança vários resultados, levando sempre em consideração os parâmetros que estão sendo avaliados, ou seja, a discriminação, a dificuldade e o acerto ao acaso. Em posse desses parâmetros, é possível fazer diversas análises, como a comparação de desempenho entre alunos, a calibração de itens e a equalização em diferentes formas de avaliação.

Alguns sistemas de avaliação usam a TRI na análise de dados, pelo fato de que, ela permite a comparabilidade dos resultados entre anos e a aplicação do teste várias vezes ao ano, isto porque, a comparação dos resultados se torna possível devido a uma escala métrica estabelecida pela TRI, definida como escala padrão de conhecimento (BRASIL, 2012).

Outra forma de se utilizar a TRI na análise de dados, é quando se deseja medir a proficiência de um aluno em determinada área de conhecimento.

4.3 Conclusão

Este capítulo apresentou como as Teorias de Medidas são utilizadas no processo de avaliação educacional, além de comentar sobre os modelos matemáticos e métodos de avaliação utilizados pela TRI. Ao final do capítulo, apresentou de que forma a TRI pode ser usada na análise de dados.

5 RECURSOS COMPUTACIONAIS QUE TRABALHAM COM A TRI

A TRI vem crescendo muito e ganhando reconhecimento de diversas instituições educacionais, e para isso ela faz uso de alguns recursos computacionais que viabilizam a sua utilização. Na Europa e nos Estados Unidos, devido a mais tempo de uso da TRI, existem vários programas específicos para análise via TRI. No entanto, no Brasil, passou-se a utilizar a TRI a pouco tempo e por isso há uma variedade bem menor de programas sendo utilizados e comercializados.

De acordo com texto explicitado, segue abaixo os programas segundo Andrade (2012).

5.1 Ferramentas

No decorrer deste capítulo, iremos comentar sobre alguns recursos computacionais que trabalham análise via TRI. Dos programas existentes no Brasil que trabalham nas análises envolvendo a TRI, existem o PARSCALE, Software R e o BILOG-MG.

Neste trabalho será usado o BILOG para calibração do teste e obtenção de resultados.

5.1.1 PARSCALE

Utilizado na análise de itens ordinais, este programa permite a análise para mais de um grupo de respondentes. Este software tem implementado os modelos de Resposta Gradual e de Créditos Parciais (ANDRADE, 2000).

5.1.2 SOFTWARE R

O Software R, não é um programa específico para análise de dados via TRI, no entanto ele é um software livre e que possui as mesmas características dos pagos, podendo então ser utilizado na análise via TRI. Ele utiliza uma linguagem de programação que permite ao usuário desenvolver suas próprias funções, de acordo com sua necessidade (ANJOS, 2010).

5.1.3 BILOG – MG

Os autores Andrade, Tavares e Vale (2000) ressaltam que o BILOG – MG, é um programa específico para análise via TRI de itens dicotômicos ou dicotomizados e ambos tem implementados os modelos unidimensionais logísticos de 1, 2 e 3 parâmetros.

O BILOG-MG permite a análise de mais de um grupo de respondentes, ele executa a análise em três etapas, chamadas de fases 1, 2 e 3, que se caracterizam pelo tipo de tarefas realizadas em cada uma delas. A fase 1 é a fase de entrada e leitura de dados e o usuário deve fornecer ao programa dois tipos de informações: A identificação de cada indivíduo com suas respectivas respostas ao teste e o gabarito. A fase 2 é a fase de calibração dos itens e são estimados os parâmetros dos itens, com seus respectivos erros-padrão. Por último, a fase 3 tem como papel fazer a estimação das habilidades dos respondentes e são estimadas as habilidades de cada um dos indivíduos, a partir dos resultados obtidos na fase anterior. O BILOG – MG já dispõe de versão para Windows e fornece gráficos que contém informações como a curva característica e as curvas de informação de cada item e do teste.

O BILOG-MG é um programa pago utilizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) para análise de itens em alguns testes, como é o caso do Exame Nacional do Ensino Médio (ENEM), do Sistema Nacional de Avaliação da Educação Básica (SAEB) e do Prova Brasil (ENEM, 2011).

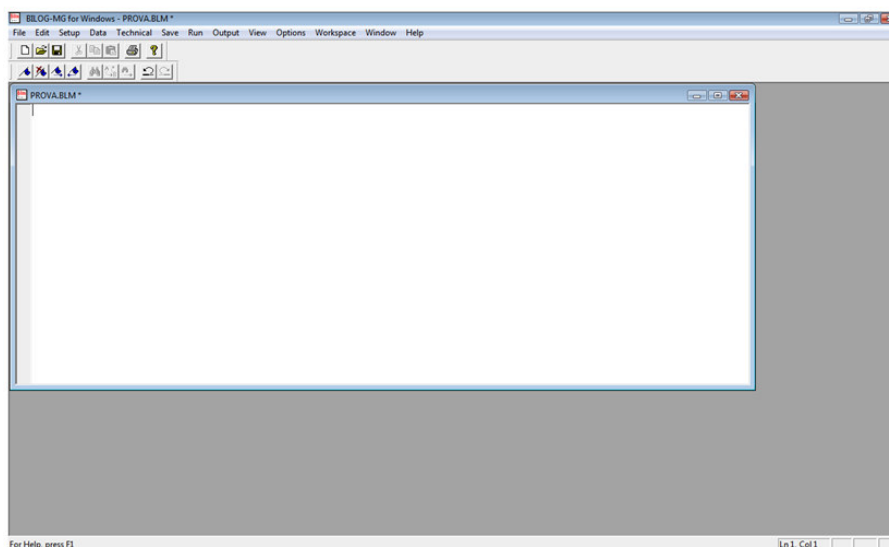


Figura 2 - BILOG - MG

5.2 Conclusão

Neste capítulo, apresentaram-se as ferramentas utilizadas para análise via TRI. Dentre as ferramentas apresentadas, as mais utilizadas são o BILOG e o BILOG-MG, conforme a situação a ser analisada.

No Brasil, são poucas as ferramentas usadas neste processo de análise envolvendo a TRI, já em outros países este tipo de análise é feita há mais tempo e por isso possui uma maior variedade de recursos computacionais.

6 MODELAGEM DO SISTEMA

Um sistema de avaliação WEB tem como principal finalidade o fato de poder ser usado por vários alunos, no mesmo momento e em diferentes locais. No trabalho proposto, a finalidade de desenvolver um sistema de avaliação WEB, deu-se pelo fato de que, o trabalho realizou um experimento em que poderá ser usado em outras pesquisas. Por ser um sistema WEB de avaliação, é possível realizar experimentos que abrangem grupos provenientes de diferentes municípios ou estados, além de poder ser implementado dentro deste sistema o CAT, que permite que alunos façam via WEB um teste adaptativo.

Para isso, na modelagem deste sistema criou-se entidades que poderão ser reutilizadas em outras áreas, com o uso ou não da TRI. As entidades criadas foram: usuário, escola, disciplina, assunto, item, alternativas, nível, Item_Prova, Prova e Questão Respondida.

Assim, deu-se início ao banco que irá possibilitar o desenvolvimento de um módulo de avaliação WEB, para isso utilizou-se uma ferramenta denominada MySQL Workbench que é uma ferramenta visual unificada para arquitetos de banco de dados, desenvolvedores e DBA's. Ela fornece modelagem de dados, desenvolvimento de SQL e ferramentas de administração abrangentes para a administração do servidor de configuração do usuário.

O módulo do sistema desenvolvido é composto de 10 tabelas evidenciadas na figura 3, que são essenciais para o seu funcionamento, possibilitando e facilitando assim a compreensão do desenvolvimento e aplicação do módulo.

A seguir, a Modelagem do Sistema de Avaliação proposto:

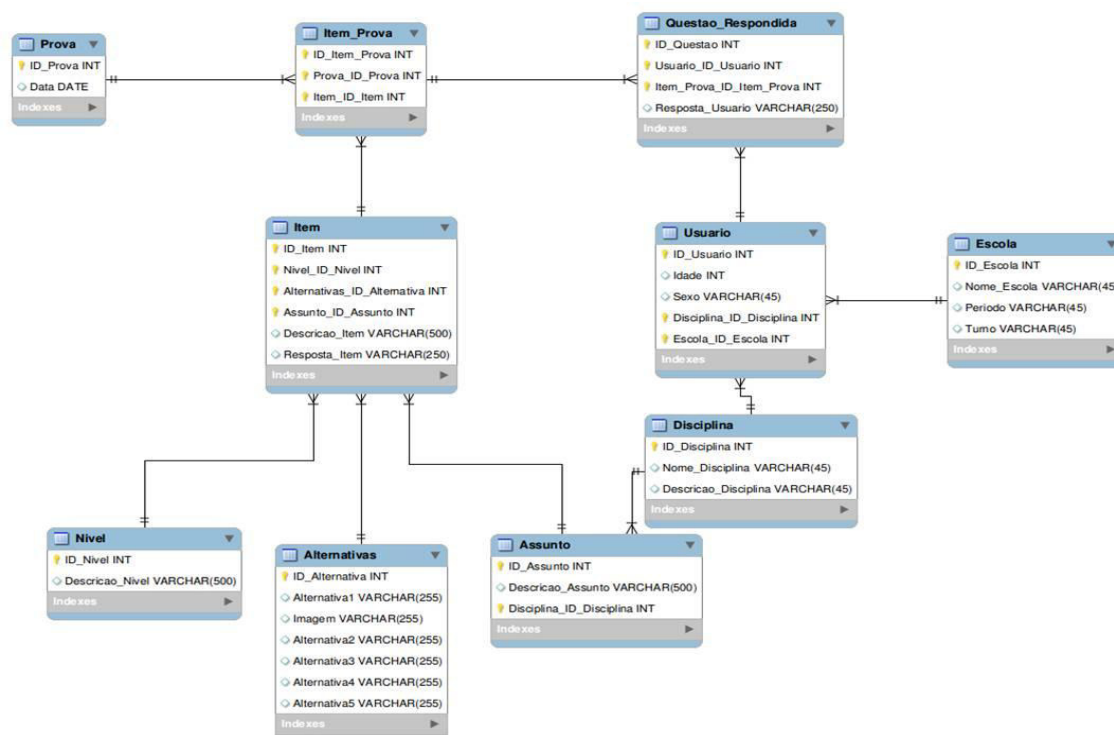


Figura 3 - Modelagem do Módulo de Avaliação proposto

Na tabela usuário é feito o cadastro do aluno, a escolha da disciplina a qual este irá se submeter ao teste e a seleção da escola a que ele pertence.

A informação da escola é guardada no campo escola e a informação da disciplina, no campo disciplina. A disciplina pode ser desmembrada por assuntos, se tratando assim de uma busca de itens por assunto. Começa-se então a compor os itens da prova.

As alternativas ficam guardadas no campo alternativas, e assim ao compor a prova no campo itens, é feita a descrição dos itens e uma seleção das alternativas que compõem aquele item. O nível faz parte do processo de modelagem, porém no que se refere à análise via TRI, a tabela nível não será utilizada, visto que, a TRI no seu processo de calibração determina o índice de dificuldade do item, mesmo assim a sua existência deve ser mantida, pois se houver a necessidade de fazer a análise pela TCM, esta tabela deverá ser utilizada.

Dessa forma, o aluno terá acesso ao Item_Prova, ou seja, neste campo a prova está pronta para ser acessada pelo aluno e respondida e as questões respondidas serão encaminhadas para o campo Questão_Respondida, possibilitando a partir das questões guardadas neste campo a análise via TRI.

O campo prova existe apenas para consulta da prova. A modelagem apresentada, deixa o banco aberto para futuros testes e implementações, assim justifica-se a existência de alguns campos.

Se este modelo for usado em outros experimentos, o campo prova, terá sua importância caso ocorra mais de uma prova e exista a necessidade de consultar a prova desejada. A tabela disciplina, se justifica pelo fato de que se futuramente houver necessidade de se fazer experimentos em outras áreas, o banco estará aberto para recebê-las.

6.1 Modelagem BPMN do Sistema de Avaliação

Vejamos o sistema de avaliação proposto, como um processo de negócio, no qual um conjunto de atividades deve ser estruturado com o objetivo de alcançar algum resultado. Como todo processo de negócio, deve-se definir uma visão e uma missão daquilo que deverá ser feito, tem-se:

- Visão - desenvolver um módulo que trabalhe com a TRI e que sirva como instrumento de coleta de dados para avaliação da migração da prova escrita, para prova no computador;
- Missão – oferecer ferramentas que facilitem e possibilitem a execução de tarefas para obtenção de resultados.

Para representar este processo de negócio foi utilizada uma linguagem gráfica, o BPMN.

O processo de negócio ilustrado na figura 4 representa o sistema de avaliação proposto:

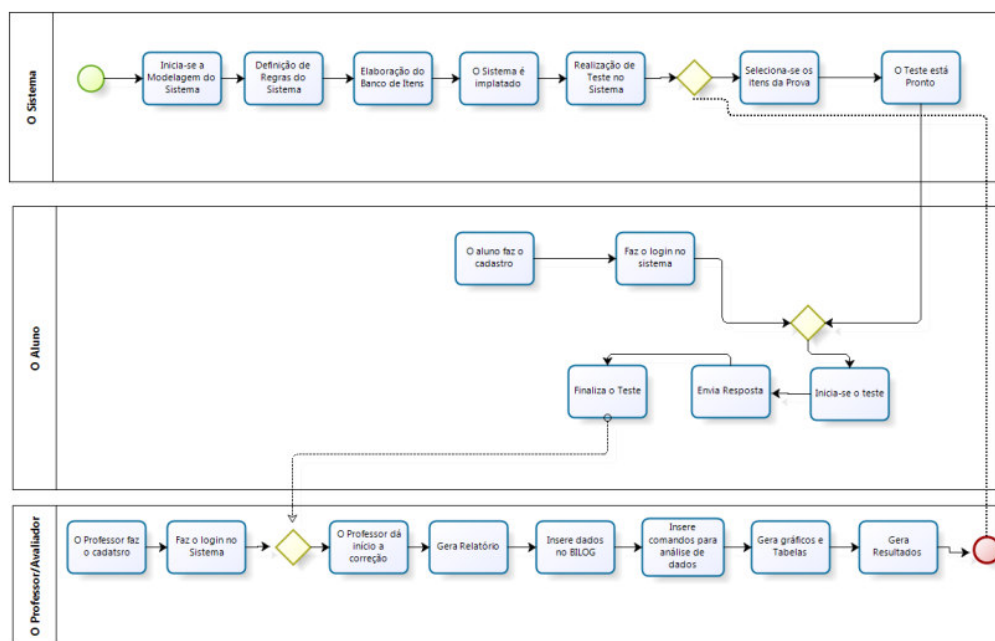


Figura 4 - Modelagem BPMN do Sistema de Avaliação

O sistema de avaliação encontra-se dividido em três partes:

- O Sistema - representa o módulo de avaliação, iniciando-se com a modelagem até a realização do teste. Se o teste realizado for aceito, ele continua, senão ele é finalizado. Quando aceito, dar-se continuidade selecionando os itens da prova;
- O Aluno - representa a interação do aluno com o teste, no entanto, para que se dê início ao teste o aluno deve estar cadastrado e com o login realizado no sistema;
- O Professor/Avaliador - representa a interação do professor com o recurso computacional usado na análise de dados via TRI, para ter acesso aos dados dos alunos, o professor/avaliador deve estar cadastrado e com o login realizado no sistema.

6.2 Diagramas de Caso de Uso do Módulo de Avaliação

O módulo de avaliação proposto já é possível de ser visualizado, no entanto, para entender de que forma foi realizado o teste, foi criado um caso de uso Fazer-Prova e Corrigir-Prova. As figuras 5 e 6 tem o objetivo de explicar

graficamente a interação entre aluno, professor e sistema e o processo da realização do teste.

O caso de Uso Fazer-Prova, no qual o aluno interage com o sistema e este emite um relatório final da prova ao professor.

Caso de Uso: FAZER – PROVA

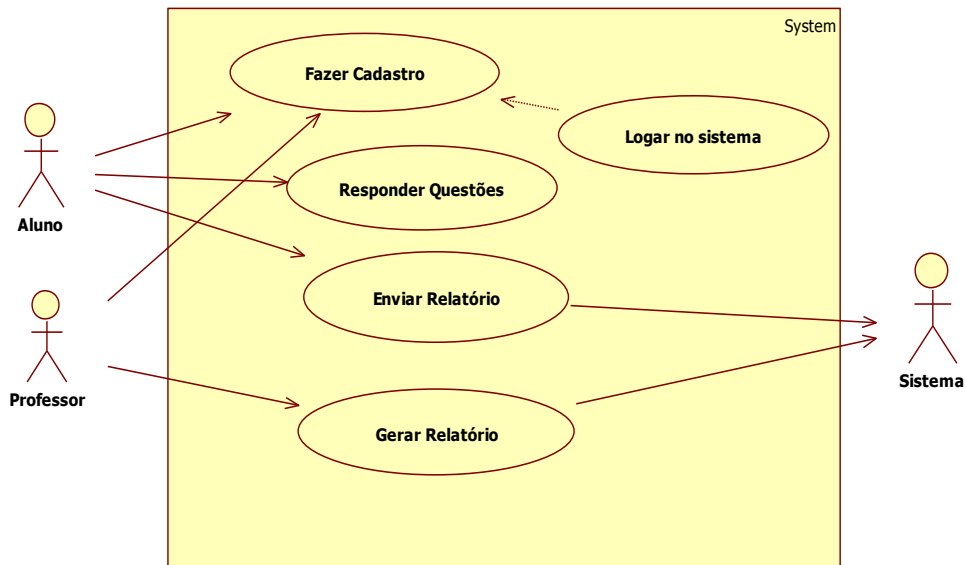


Figura 5 – Diagrama de Caso de Uso Fazer - Prova

Caso de uso: Fazer- Prova

Descrição: Este caso de uso tem por objetivo apresentar o fluxo de tarefas realizadas entre os atores para a realização do teste.

Atores: Aluno;

Professor;

Sistema.

Pré- Condição: O aluno e o professor devem estar cadastrados no sistema.

Pós- Condição: Relatório do aluno ficará mantido no sistema.

Fluxo Principal:

- O caso de uso é iniciado quando o aluno acessa o sistema e clica no ícone “você ainda não tem o login?”;
- O sistema apresenta a interface de cadastro de usuário, neste item o aluno deverá preencher os campos solicitados;
- O aluno faz o login no sistema;

- O aluno agora possui acesso às questões e começa a resolvê-las
- Ao término de cada questão, o aluno deverá clicar em “Enviar resposta”;
- Ao final da última questão, o aluno recebe a seguinte mensagem: Parabéns!

Você terminou nosso teste!;

- Neste momento, o aluno tem seu relatório de resposta consolidado;
- Neste momento, o professor faz o seu login no sistema;
- O professor tem a acesso ao relatório final do teste de cada aluno.

Fluxo Alternativo:

- Cadastro realizado com sucesso: O usuário recebe esta mensagem ao concluir o cadastro no sistema;
- Campos com * são obrigatórios: O usuário recebe esta mensagem se não preencher o campo obrigatório;
- Nome ou senha inválida: O usuário deverá receber esta mensagem se colocar um caractere inválido.

O caso de uso Corrigir-Prova, em que o avaliador interage com o Recurso Computacional, ferramenta utilizada na análise de dados via TRI.

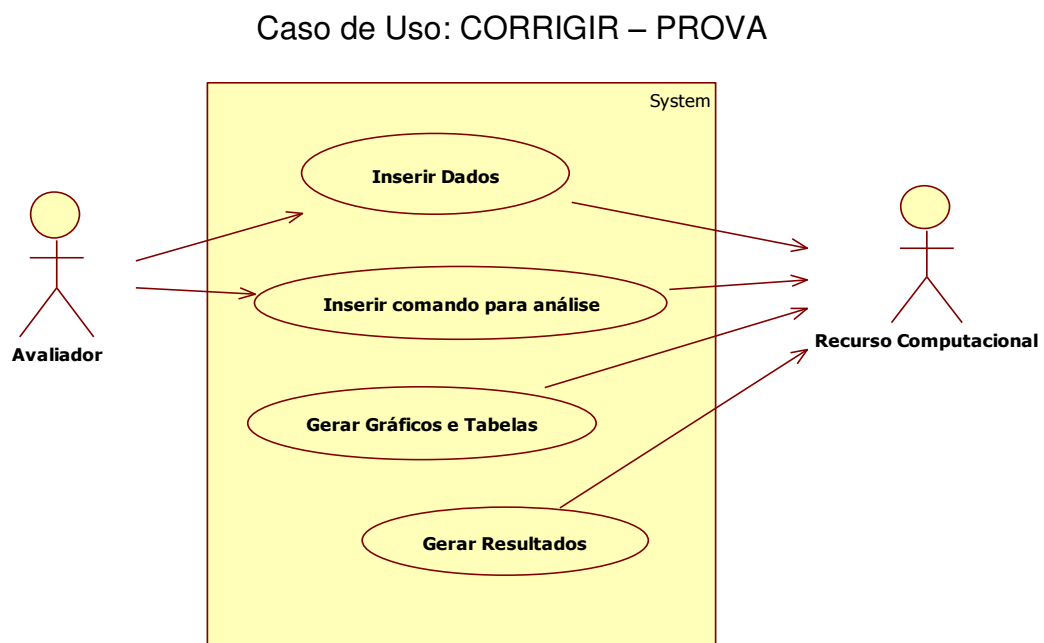


Figura 6 – Diagrama Caso de Uso Corrigir - Prova

Caso de uso: Corrigir- Prova

Descrição: Este caso de uso tem por objetivo apresentar o fluxo de tarefas realizadas entre o avaliador e o sistema na análise de dados.

Atores: Avaliador e sistema.

Pré- Condição: O avaliador deve ter acesso aos resultados individuais de cada aluno.

Pós- Condição: Os itens deverão ser calibrados.

Fluxo Principal:

- O caso de uso é iniciado quando o avaliador insere os dados do aluno no sistema, para avaliação;
- O avaliador inicia a análise de dados inserindo comandos necessários para o mesmo;
- O Sistema gera gráficos e tabelas, conforme solicitação do avaliador;
- O Sistema gera resultados.

6.3 Conclusão

Este capítulo é essencial na compreensão do processo de avaliação educacional que foi desenvolvido e que será aplicado. Pois nele foi apresentado a modelagem do banco, que compreende o módulo de avaliação, e a modelagem BPMN do Sistema de Avaliação, que compreende o processo como um todo, além de apresentar ainda os casos de uso: Fazer – Prova e Corrigir – Prova, que mostram a interação entre aluno, professor e sistema.

Este capítulo justifica ainda o fato de se trabalhar com um sistema de avaliação WEB. No capítulo a seguir, serão apresentadas algumas avaliações existentes no Brasil e no Mundo que fazem uso da TRI na análise de dados, seguida da apresentação da metodologia e dos resultados do sistema de avaliação proposto.

7 APLICAÇÃO DA TRI NO PROCESSO DE AVALIAÇÃO

As Teorias de Medidas são utilizadas nos diversos processos de avaliação existentes no Brasil e no mundo. A maioria dos processos utiliza a Teoria Clássica de Medidas, porém existem vários processos de avaliação que já fazem uso da TRI como método de Avaliação. Assim podemos apresentar o Sistema de Avaliação da Educação Básica (SAEB), em que o uso da TRI a este processo teve início em 1995 e, sendo implementado posteriormente, no ENCCEJA, Prova Brasil e ENEM. Se tratando de avaliação em âmbito internacional, a TRI vem sendo utilizada por vários países: Estados Unidos, França, Holanda, Coreia do Sul, China, sem falar nos países participantes do PISA (BRASIL, 2012).

Um dos grandes exemplos de avaliação utilizando a TRI é o SAT, um exame educacional padronizado dos EUA aplicado a estudantes do Ensino Médio, e o exame de proficiência em Língua Inglesa (TOEFL), este exame surgiu em 1964 e é amplamente utilizado em todo o mundo. Desde o ano de sua origem, este exame já avaliou mais de 25 milhões de alunos e tem sido administrado por mais de 4.500 centros em 165 países do mundo. No TOEFL, os alunos marcam o horário em um dos centros credenciados e podem realizá-lo varias vezes ao ano. A prova é adaptativa, realizada no computador, e cada candidato responde a um conjunto de itens (questões) diferentes. Apesar de em cada aplicação o candidato receber uma prova distinta, todos os resultados são comparáveis e considerados isonômicos (BRASIL, 2012).

No processo de avaliação desenvolvido para experimento desta dissertação, foram realizados dois tipos de testes, sendo um seguindo os moldes tradicionais no papel e caneta, e o outro realizado utilizando recursos computacionais, ou seja, feito no computador. Este processo também fez uso da TRI na análise dos dados, e no decorrer desse capítulo será apresentado de que forma o processo foi realizado e como a TRI foi aplicada ao mesmo.

7.1 Metodologia dos Testes

Os testes foram realizados nos dias 18, 19 de outubro nos turnos matutino e vespertino. Participaram do teste os alunos do 5º ano do ensino fundamental das seguintes escolas: Escola Municipal Raimundo Chaves, Escola Municipal Rosália Freire, Escola Municipal Rosário Nina e o Colégio Universitário- COLUN. Cada grupo de cada escola era separado de forma aleatória, de maneira que 50% da turma fizesse a prova escrita e os outros 50% fizesse o teste no computador.

A prova no computador teve uma aceitação maior, pois os alunos demonstravam um maior interesse na realização do teste, no entanto houve certa rejeição por parte dos alunos que nunca tiveram contato com a máquina, dificultando assim que, o aluno voltasse à questão anterior quando sentisse necessidade.

A prova escrita é amplamente conhecida devido a sua tradicional metodologia que o aluno recebe a prova, preenche o campo com os dados pessoais e após responder, entrega para o professor responsável. Foi dessa forma que se realizou o teste escrito. No entanto foi inserido no software de avaliação, um código referente a cada aluno, com as repostas referentes ao mesmo, assim a correção da prova escrita foi efetuada da mesma forma da correção da prova feita no computador.

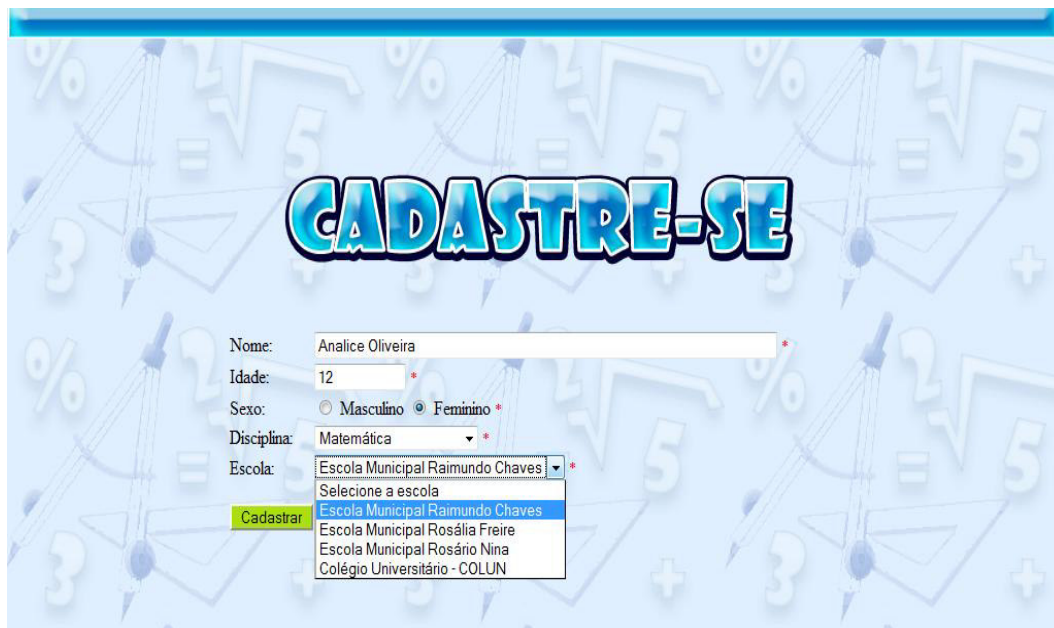
7.2 Aplicação do Teste

Conforme apresentado, o módulo WEB de avaliação já está pronto, e, portanto já pode ser aplicado aos alunos do 5º ano. Os estudantes que se encontram neste grupo possuem um conhecimento sobre a matemática aplicada no teste, e por isso estão preparados para fazer a prova.

O aluno deve ter a consciência de que a avaliação implica um fundamentado conhecimento sobre o que interrogamos e atribuição de significados aos fatos, dados e informações que colhemos (SGUISSARDI, 1997).

O teste aplicado compõe-se de 30 itens, onde 8 deles são itens da escala SAEB, isto porque, para que a TRI calibre estes itens é importante que existam itens envolvidos no teste e que já foram calibrados anteriormente.

Para ter acesso a avaliação, o aluno deverá se cadastrar e fazer o login no sistema, conforme figuras 7 e 8 abaixo:



A tela de cadastro apresenta o título "CADASTRE-SE" em letras grandes e coloridas. Abaixo, há campos para preenchimento de dados pessoais e acadêmicos. O campo "Nome" contém "Analice Oliveira". O campo "Idade" contém "12". O campo "Sexo" tem duas opções: "Masculino" (desselecionada) e "Feminino" (selecionada). O campo "Disciplina" é um menu suspenso com "Matemática" selecionada. O campo "Escola" é um menu suspenso com "Escola Municipal Raimundo Chaves" selecionada. Abaixo dos campos, há um botão verde "Cadastrar".

Nome: Analice Oliveira *

Idade: 12 *

Sexo: ☐ Masculino ☒ Feminino *

Disciplina: Matemática *

Escola: Escola Municipal Raimundo Chaves *

Selecionar a escola

Escola Municipal Raimundo Chaves

Escola Municipal Rosália Freire

Escola Municipal Rosário Nina

Colégio Universitário - COLUN

Cadastrar

Figura 7 - Módulo de Avaliação - Tela Cadastro



A tela de login apresenta o título "DEZ E DEZ" em letras grandes e coloridas. Abaixo, há campos para preenchimento de dados pessoais e acadêmicos. O campo "Nome" contém "Analice Oliveira". O campo "Senha" contém "....". Abaixo dos campos, há dois botões: "Entrar" (verde) e "Limpar" (verde). Abaixo dos botões, há uma mensagem: "Campos com * são obrigatórios!". No canto inferior direito, há um ícone de documento e o texto "Você ainda não tem login?".

Nome: Analice Oliveira *

Senha: *

Entrar Limpar

Campos com * são obrigatórios!

Você ainda não tem login?

Figura 8 - Módulo de Avaliação: Fazer Login

Quando o aluno preenche o cadastro determinando o sexo, ele é direcionado para uma tela login, efetuado o login com nome e senha ele inicia o teste e a prova apresentada é personalizada com o seu nome e com característica

feminina ou masculina. Essa foi uma forma simples e eficaz de prender a atenção do aluno naquilo que ele está respondendo.

As figuras 9, 10 e 11 apresentadas a seguir, exemplificam três itens iguais que foram aplicados na prova de maneiras diferentes:

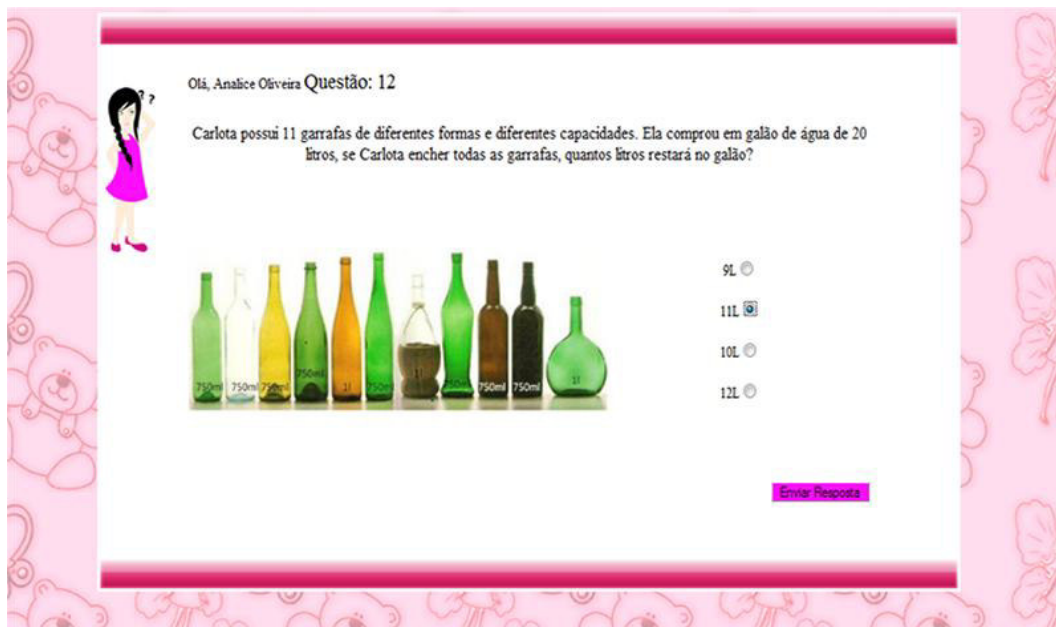


Figura 9 - Item 12 do Módulo de Avaliação aplicado no computador para meninas

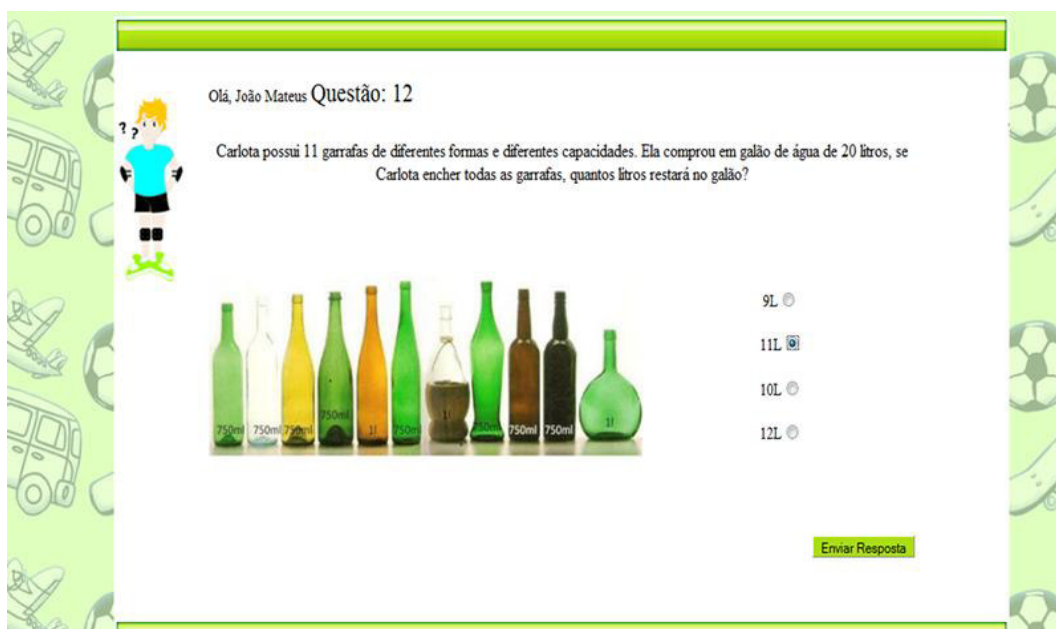
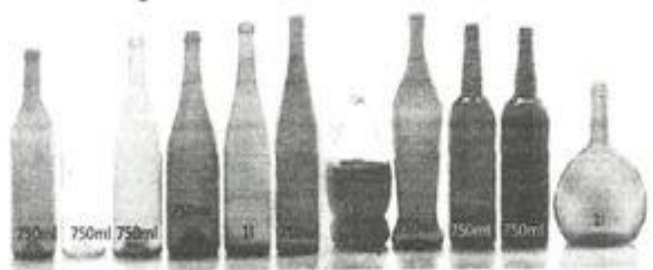


Figura 10 - Item 12 do Módulo de Avaliação aplicado no computador para meninos

12. Carlota possui 11 garrafas de diferentes formas e diferentes capacidades. Ela comprou em galão de água de 20 litros, se Carlota encher todas as garrafas, quantos litros restará no galão?



a) 9L
☒ b) 11L
c) 10L
d) 12L

Figura 11- Item 12 do Módulo de Avaliação aplicado no papel para meninos e meninas

O item acima quer extrair do aluno um conhecimento sobre as unidades de medidas e as quatro operações, de forma que o aluno avaliado deve interpretar a questão e combinar as operações necessárias na resolução do problema.

Todos os 30 itens foram apresentados em três formas, no entanto ao final do teste feito com papel e lápis, o sistema não emite relatório de avaliação, enquanto que, ao final do teste realizado no computador, o sistema emite um relatório para o avaliador, conforme figura 12:

MASCULINO				
ID	Aluno	Acertos	Erros	Data e Hora
185	thaisson	11	19	2012-10-25 13:28:54
186	Leonardo	15	15	2012-10-26 09:48:16
188	Jadson	13	17	2012-10-24 19:14:01
194	pedro lucas	11	19	2012-10-25 13:50:12
195	jefferson philipe	13	17	2012-10-25 13:58:27
196	deice	9	21	2012-10-25 14:15:00
197	clerilson	8	22	2012-10-25 14:15:01
199	lucas	9	21	2012-10-25 14:15:00
200	Lauanderson	7	23	2012-10-25 14:10:15
202	Jhonata Henrique	14	16	2012-10-25 14:53:06
213	ellen	20	10	2012-10-26 10:07:46
214	Luis Vitor	23	7	2012-10-26 10:13:33
215	Bruno Silva	19	11	2012-10-26 10:23:10
216	isaac	22	8	2012-10-26 10:28:26
218	carlos eduardo	17	13	2012-10-26 10:41:49
220	matheus ferreiro	22	8	2012-10-26 10:58:00
221	Juapablo costa	17	13	2012-10-26 11:07:25
FEMININO				
ID	Aluno	Acertos	Erros	Data e Hora
204	yanka	14	16	2012-10-26 08:35:37
205	wendy	13	17	2012-10-26 09:22:54
206	Yasmin Araújo	16	14	2012-10-26 09:33:55
207	Suene	15	15	2012-10-26 09:28:26
210	Layla Bianca	4	3	2012-10-25 15:04:48
212	Samanta Pires	21	9	2012-10-26 09:58:42
217	Luciele	18	12	2012-10-26 10:34:43
219	angela	19	11	2012-10-26 10:49:32
222	ana lucia	19	11	2012-10-26 11:13:18
224	Iussir cris	19	11	2012-10-26 11:23:02
225	jaila barros	21	9	2012-10-26 11:27:41
226	mibia de cassia	17	13	2012-10-26 11:33:14
227	cassandra mota pereira	15	15	2012-10-26 11:40:09
228	Thalita cristina soares ribeiro	15	15	2012-10-26 11:50:32
229	Anaíse Oliveira	24	6	2012-10-29 09:52:50

Figura 12 - Relatório Final de Avaliação

7.3 Testes e Resultados do Sistema de Avaliação aplicados

Após a apresentação de modelagens e diagramas que representam o processo de avaliação em pleno funcionamento, neste item serão apresentados os

resultados encontrados com a avaliação in loco do módulo de avaliação desenvolvido aos alunos do 5º ano do ensino fundamental.

Na escolha de itens para o teste, o SAEB ofereceu alguns itens já calibrados para que dentre eles, fossem escolhidos 8 itens para compor o teste. Assim é possível calibrar alguns itens e manter o padrão dos outros. Observe a figura abaixo adaptada pelo autor deste trabalho.

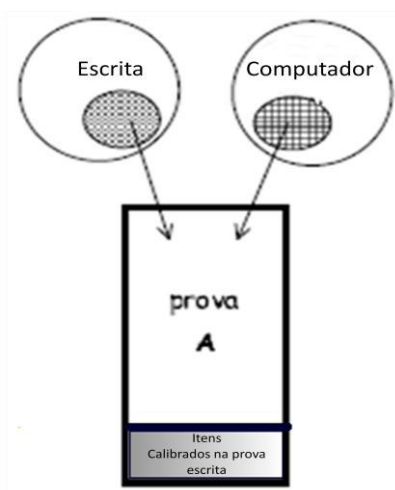


Figura 13 - Representação gráfica da aplicação da prova “Dois grupos fazendo uma única prova” (ANDRADE, 2000).

Na análise feita utilizando a TRI, a estimação dos parâmetros dos itens foi feita por máxima verossimilhança marginal e a estimação das proficiências utilizou-se o Método de Estimação Bayesiano EAP (verificar no capítulo 4). Dessa forma, o procedimento disponível no programa utilizado para fixar apenas parte dos itens, é o uso de distribuições a priori convenientes para os parâmetros desse item. Para os itens novos que se deseja calibrar, utiliza-se as prioris padrão sugeridas pelo programa. Já para os outros itens definimos prioris cujas médias são os próprios valores dos parâmetros que desejamos fixar e cujo desvio é tão pequeno que a distribuição torna-se praticamente degenerada naquele ponto (ANDRADE, 2000).

No trabalho aqui apresentado foi utilizado o BILOG, na análise em que dois grupos são submetidos a uma mesma prova, neste caso a equalização é via itens comuns, sendo necessário o uso de recursos computacionais para análise via TRI. Utilizando o BILOG, o teste foi calibrado separadamente e depois foi feita uma equalização a posteriori, conforme foi descrito no capítulo 4.

A seguir serão apresentadas as tabelas 2 e 3 referente aos resultados das análises de dados feitas via TRI no BILOG. Os respondentes são amostras

aleatórias da mesma população, podemos dizer assim que o uso da escala 0,1 nas duas calibrações representam os mesmos valores, portanto os resultados estarão na mesma escala, tendo a equalização feita.

Tabela 2 - Calibração na escala 0,1 da prova escrita

Escrita: calibração na escala (0,1)						
Item	a	ep(a)	b	ep(b)	c	ep(c)
ESC01	1,112	0,343	0,445	0,341	0,209	0,081
ESC02	não calibrou					
ESC03	0,986	0,279	0,493	0,381	0,208	0,089
ESC04	não calibrou					
ESC05	0,759	0,276	1,991	0,740	0,202	0,071
ESC06	1,052	0,327	2,236	0,630	0,203	0,090
ESC07	0,902	0,353	2,572	0,857	0,174	0,053
ESC08	0,985	0,353	2,100	0,620	0,130	0,050
ESC09	0,984	0,347	2,090	0,611	0,120	0,049
ESC10	1,603	0,423	0,193	0,233	0,181	0,075
ESC11	1,074	0,292	0,852	0,362	0,195	0,086
ESC12	não calibrou					
ESC13	0,570	0,178	0,016	0,613	0,219	0,093
ESC14	1,465	0,376	0,748	0,268	0,180	0,079
ESC15	1,199	0,338	0,132	0,297	0,186	0,078
ESC16	1,049	0,312	0,231	0,350	0,206	0,083
ESC17	não calibrou					
ESC18	1,692	0,449	0,691	0,256	0,201	0,084
ESC19	2,044	0,564	0,200	0,182	0,170	0,063
ESC20	0,927	0,294	0,591	0,414	0,214	0,084
ESC21	1,142	0,369	1,029	0,353	0,172	0,066
ESC22	não calibrou					
ESC23	1,199	0,448	2,002	0,526	0,128	0,046
ESC24	não calibrou					
ESC25	1,054	0,428	2,285	0,720	0,206	0,055
ESC26	não calibrou					
ESC27	não calibrou					
ESC28	0,865	0,313	1,938	0,650	0,178	0,064
ESC29	1,064	0,423	1,965	0,613	0,225	0,061
ESC30	1,112	0,419	1,981	0,556	0,152	0,052
média	1,129	0,359	0,743	0,481	0,184	0,071

Observe na tabela 2, que dentre as 30 questões que se submeteram ao processo de calibração, oito delas não foram calibradas.

A prova escrita (ver anexo) apresentou algumas desvantagens em relação à prova feita no computador, nela algumas questões tiveram as figuras prejudicadas pela impressão, além de que os alunos se mostraram mais interessados na prova feita no computador. Nela o aluno sentia-se interagindo com a prova, no entanto no papel, além de ser comum ao seu convívio escolar, ele se sente pressionado e aterrorizado com a situação de se estar fazendo uma prova, principalmente de matemática.

Na tabela 2 verifique que o item 2 não foi calibrado, observe a figura referente ao Item:

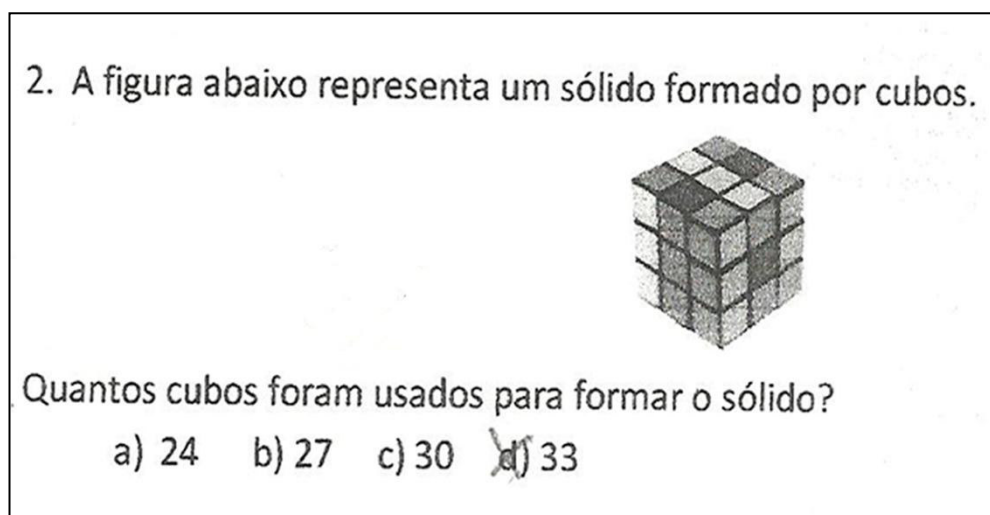


Figura 14 - Item 02 do Módulo de Avaliação aplicado no papel para meninos e meninas

Isso se justifica pelo fato de que no computador a figura é mais clara e visível, ou seja, mais fácil de ser interpretada, sendo assim a questão é prejudicada quando aplicada no papel.

Existem outros itens que também não foram calibrados, e que possuem diferentes justificativas para tal, são eles os itens: 04, 12, 17, 22, 24, 26, 27.

O item 4 foi prejudicado pelo desenho, que confundiu o raciocínio, ou seja, a questão não foi bem elaborada.

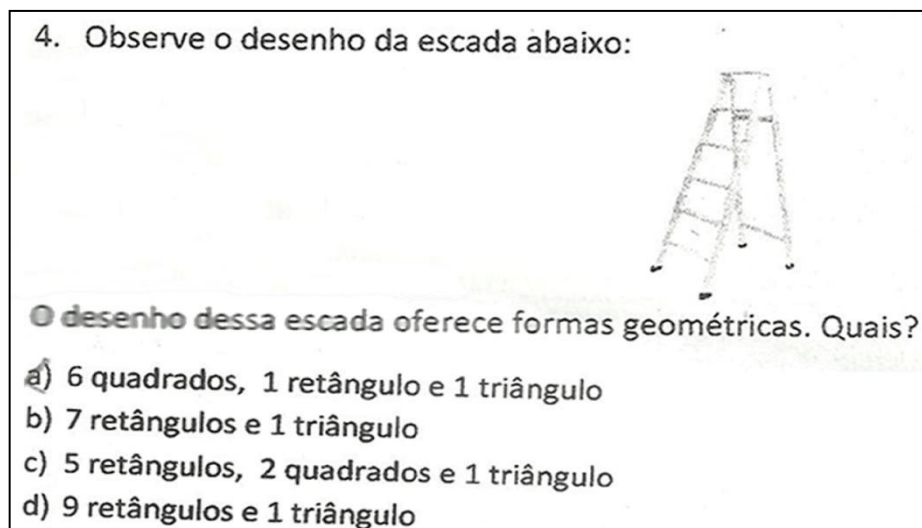


Figura 15 - Item 04 do Módulo de Avaliação aplicado no papel para meninos e meninas

Note que o desenho da escada não é esclarecedor, podendo haver várias respostas para a questão: “Quantas formas geométricas é possível extrair da figura?”, neste caso, o número de possibilidades apresentadas nas opções era inferior ao número real de possíveis formas geométricas.

O item 12, conforme apresenta o item 7.1 da dissertação, teve a figura referente a questão prejudicada devido a má qualidade da impressão, da mesma forma que aconteceu no item 2.

O item 17 apresenta a figura de um relógio sem números, e isso causou uma reclamação entre os estudantes que não sabiam verificar horas em relógio desse modelo.

Os itens 22 e 24 são itens já calibrados anteriormente e que não foram calibrados nessa avaliação. A justificativa encontrada para tal é a falta de interpretação das questões, o que causou um alto índice de respostas erradas e em branco. (Figuras 16 e 17).

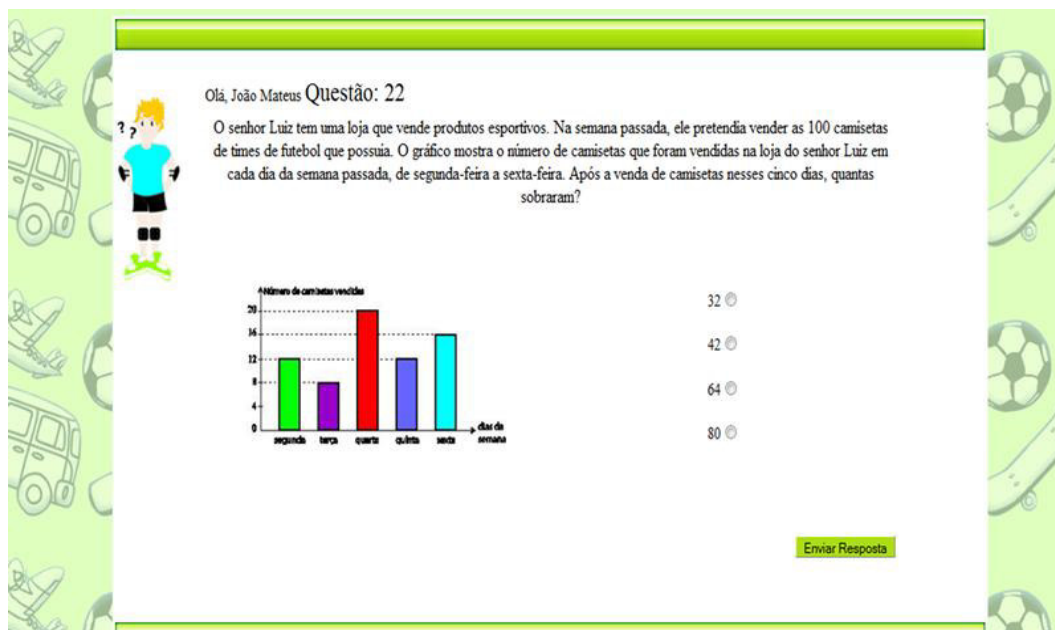


Figura 16 - Item 22 do Módulo de Avaliação aplicado no computador para meninos



Figura 17 - Item 24 do Módulo de Avaliação aplicado no computador para meninos

O item 26 envolvia uma regra de três simples na resolução das questões, no entanto isso causou uma reclamação entre estudantes que justificaram não ter conhecimento sobre este tipo de regra matemática, isto ocorreu devido ao fato de a escola estar atrasada em relação ao conteúdo, pois conforme plano de ensino, este assunto já deveria ter sido tratado com os alunos.

O item 27, assim como os itens 22 e 24, envolvia uma interpretação da situação do problema e operações matemáticas, no entanto, os alunos sentiram dificuldade em interpretá-lo, causando assim um índice elevado de respostas erradas e em branco.

As demais questões foram calibradas.

Abaixo será apresentada a tabela resultante da calibração feita da prova no computador:

Tabela 3 - Calibração na escala 0,1 da prova no computador

Computador: calibração na escala (0,1)						
Item	a	ep(a)	b	ep(b)	C	ep(c)
COM01	não calibrou					
COM02	0,884	0,265	-0,042	0,403	0,196	0,084
COM03	não calibrou					
COM04	não calibrou					
COM05	0,976	0,387	2,476	0,810	0,141	0,049
COM06	não calibrou					
COM07	não calibrou					
COM08	0,793	0,290	1,550	0,647	0,220	0,078
COM09	0,932	0,358	2,187	0,724	0,161	0,057
COM10	não calibrou					
COM11	não calibrou					
COM12	não calibrou					
COM13	0,874	0,278	0,261	0,437	0,216	0,088
COM14	não calibrou					
COM15	1,723	0,456	-0,384	0,224	0,165	0,071
COM16	não calibrou					
COM17	não calibrou					
COM18	1,634	0,466	-1,336	0,329	0,191	0,084
COM19	não calibrou					
COM20	não calibrou					
COM21	0,901	0,364	2,916	1,021	0,131	0,046
COM22	não calibrou					
COM23	não calibrou					
COM24	não calibrou					
COM25	não calibrou					
COM26	não calibrou					
COM27	não calibrou					
COM28	não calibrou					
COM29	não calibrou					
COM30	não calibrou					
média	1,090	0,358	0,954	0,574	0,178	0,070

Na prova feita no computador, verificou-se um índice muito maior de itens não calibrados, do que na prova feita no papel, este resultado foi inesperado às nossas expectativas. Apenas os itens 02, 05, 08, 09, 13, 15, 18 e 21 foram calibrados.

Os estudantes na sua grande maioria nunca tiveram aula de informática, ainda assim ficaram extasiados com a prova no computador, alguns já haviam tido contato com essa tecnologia, no entanto nunca haviam feito pesquisas ou usado o computador como ferramenta complementar do ensino-aprendizagem, o pouco contato que tem com este recurso é para uso em jogos e redes sociais.

Na foto ilustrada na figura 18 é possível verificar a euforia dos alunos em laboratório, foram preciso três professores para fiscalizar e controlar a situação.



Figura 18 - Aplicação do teste em laboratório para meninos e meninas

Levando em consideração que os alunos poucas vezes tiveram contato com o uso dessa ferramenta, atribuo a isto o fato de ter um resultado tão inesperado na prova feita no computador, já que, teoricamente a prova no computador é mais interativa e com resultados positivos.

O item 2 conforme apresentado na tabela 2 e 3, não foi calibrado na prova escrita e foi calibrado na prova no computador, justifica-se pelo fato de que, no computador a imagem é mais esclarecedora, visível e dá a impressão de interatividade.

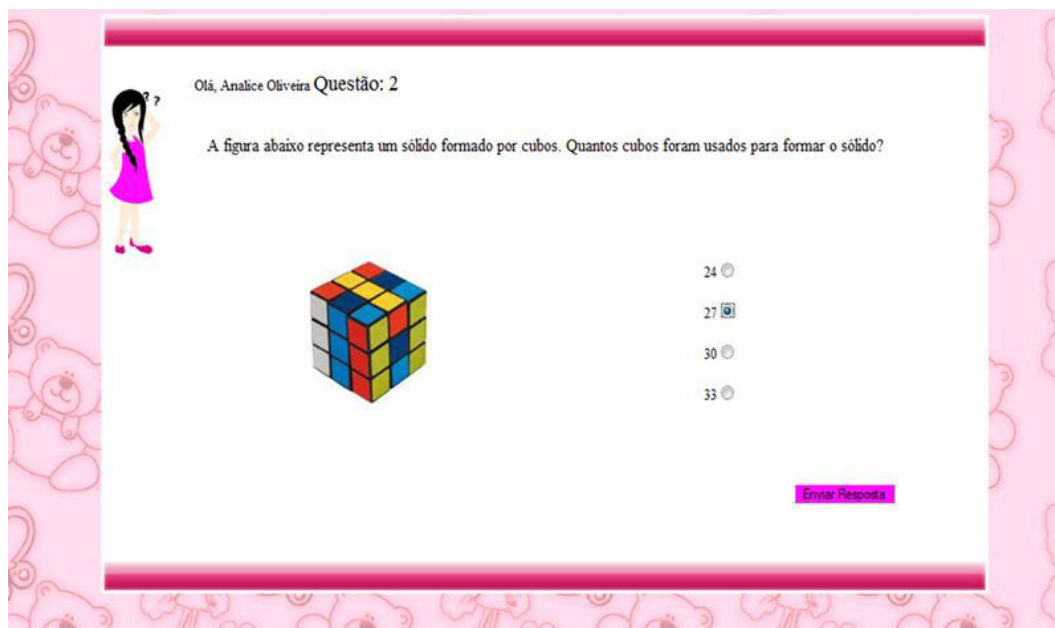


Figura 19 - Item 02 do Módulo de Avaliação aplicado no computador para meninas

O item 4 não foi calibrado em nenhum modelo de aplicação de prova, isto porque a imagem seja ela no papel ou no computador, não responde à pergunta em questão. Neste caso o item seria anulado.

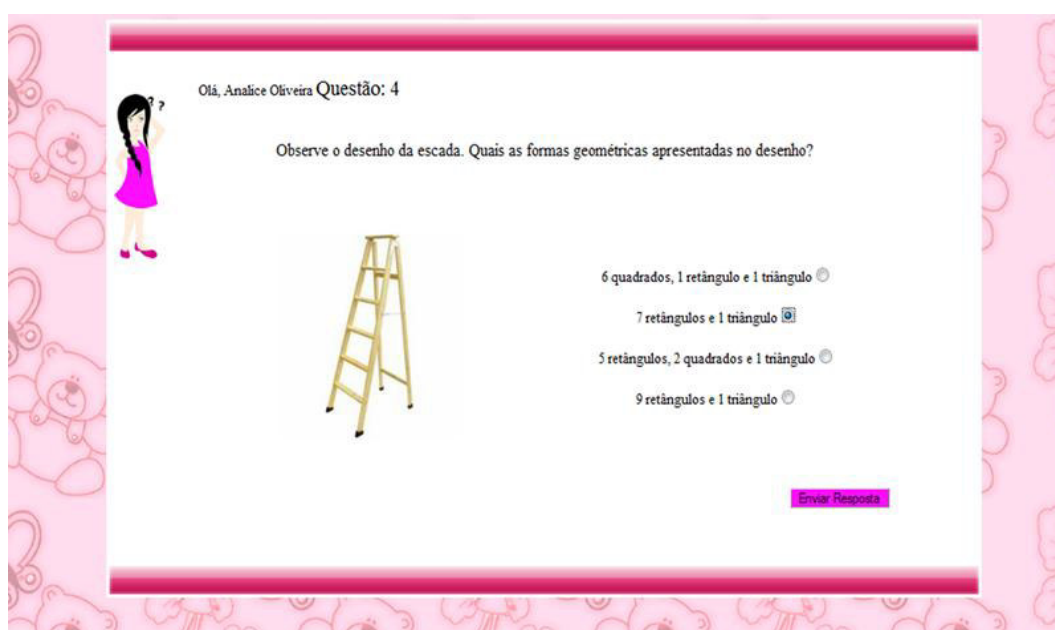


Figura 20 - Item 04 do Módulo de Avaliação aplicado no computador para meninas

O item 12 conforme apresenta o item 7.1 da dissertação, não foi calibrado em nenhum modelo de prova, isto se deve ao fato de que, a figura não foi

esclarecedora em nenhuma situação, na prova escrita à quantidade de ml/litros apresentada nas garrafas, não foi clara em pelo menos quatro modelos de garrafas, assim também aconteceu na prova feita no computador, em que em pelo menos um modelo de garrafa, a quantidade de ml/litro não foi clara. Assim essa questão foi prejudicada não sendo calibrada em nenhum modelo de avaliação.

Os demais itens que foram calibrados no computador também foram calibrados no papel.

Os itens 05, 09, 13 e 21 tiveram o parâmetro **(b)** da prova no computador, maior que o parâmetro **(b)** da prova no papel, isso devido ao fato de que, essas questões são mais difíceis no computador do que no papel.

Os itens 08, 15 e 18 tiveram o parâmetro **(b)** da prova no computador, menor que o parâmetro **(b)** da prova no papel, isso indica que essas questões são mais fáceis no computador do que no papel.

Tabela 4 - Comparação entre itens calibrados na prova escrita e no computador

Comparação entre itens calibrados		
a_E/a_C	b_E - b_C	c_E - c_C
0,777	-0,485	0,061
1,242	0,549	-0,090
1,056	-0,098	-0,041
0,653	-0,245	0,003
0,696	0,516	0,021
1,035	0,645	0,010
1,267	-1,888	0,041
0,961	-0,143	0,001

No processo de comparação entre os tipos de avaliações aplicadas apresentada na tabela 4, a comparação entre as estimativas dos parâmetros dos itens não obtiveram um padrão nos quais os b's em uma prova apresentassem resultados maiores que os b's na outra prova, no entanto neste processo de comparação isto não ocorreu.

7.4 Análise dos Resultados

Logo, a TRI na sua análise em relação ao impacto causado na migração da prova escrita para a prova no computador, mostrou que apesar das vantagens que o computador tem quando comparado à prova escrita, a aplicação do teste no computador deve ser bem planejada, e os grupos envolvidos no teste devem ter conhecimento e saber manipular bem a ferramenta, pois apesar do grande interesse dos estudantes em fazer um teste no computador, o poder de distração que essa ferramenta pode causar é muito maior do que na prova escrita.

Conforme a análise feita na figura abaixo, é possível perceber as vantagens e desvantagens de se fazer a prova no papel ou no computador.

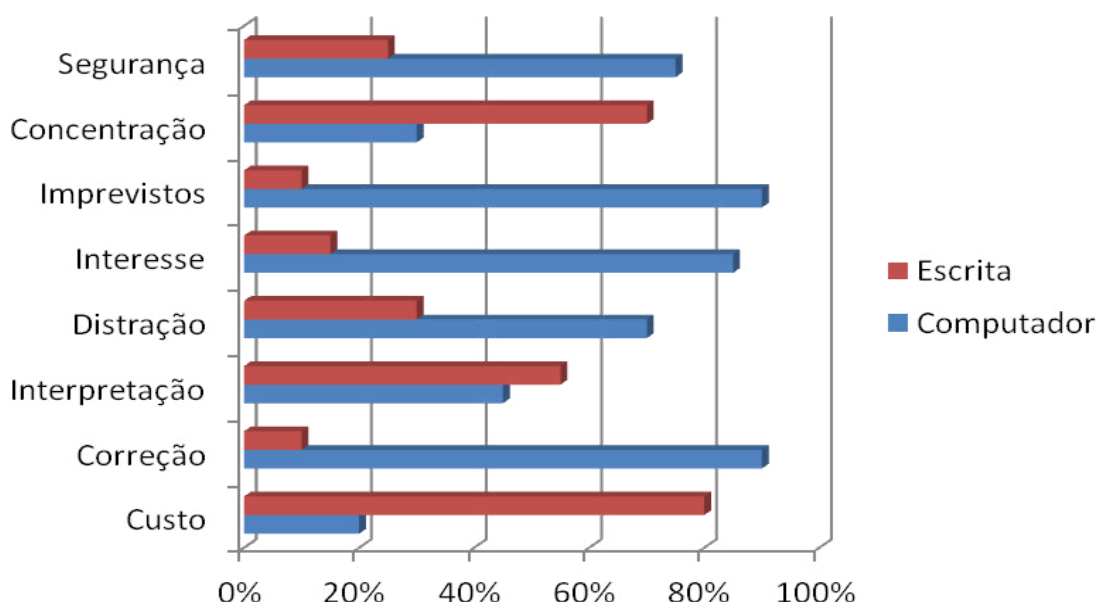


Figura 21 – Gráfico Comparativo entre a prova escrita e no computador

Os resultados encontrados na análise feita pela TRI mostram que a prova no computador teve apenas oito itens calibrados dos trinta itens existente (Ver tabela 3). Conforme a análise feita durante o teste nota-se que, a concentração dos alunos foi comprometida na aplicação da prova no computador, além de que, das quatro escolas participantes do processo, três delas fazem parte da rede municipal de ensino, que por sua vez não oferece infraestrutura e nem tão pouco aulas de informática para seus alunos, o que prejudicou o desempenho dos mesmos durante a prova. Observou-se por meio de indagações feitas aos coordenadores e alunos, que o ensino referente à informática está sendo prejudicado devido a falta de

professores capacitados na área e a falta de infra-estrutura oferecida pela instituição, assim os alunos dizem sentir-se prejudicados pela falta de contato com a tecnologia.

Os alunos do 5º ano fundamental das escolas municipais já foram submetidos a Prova Brasil, que é uma prova escrita, que possui o mesmo padrão da prova aplicada nessa pesquisa, e que também tem a análise feita pela TRI, ou seja, a prova escrita possuía um padrão comum e mais próximo da realidade dos alunos.

No entanto, na análise comparativa entre os dois modelos de prova verificou-se que a segurança dos itens é maior no computador, pois durante a confecção das provas escritas os itens são expostos nas mãos de muitos, no computador isso não ocorre; a concentração dos alunos está relacionada à distração dos mesmos, ou seja, na prova escrita os alunos tendem a ter um maior poder de concentração, por ser algo comum e cotidiano na vida escolar, enquanto que na prova no computador os alunos se distraem mais fácil, por ser algo diferente e com mais de recursos, que vão além do modelo de aprendizagem educacional e para alguns é um momento único na vida escolar.

Por isso justifica o comportamento interessado dos alunos em relação à prova no computador.

Imprevistos acontecem em qualquer lugar, ambiente ou natureza, no entanto quando se trata de prova, os imprevistos podem ser maiores no computador, apresentam-se em diversas formas: Um problema na máquina (hardware ou software), na internet (em casos de prova WEB), na energia e até problemas pessoais, quando o estudante ou por falta de conhecimento ou por alguma limitação física, não consegue manipular a ferramenta.

Os alunos apresentaram dificuldades quando se trata de interpretação dos problemas, sendo o índice maior na prova feita no computador, em que o fator “concentração” é menor.

Houve questão ainda, em que os alunos foram prejudicados devido a atrasos na grade escolar.

No fator correção, a prova no computador é mais ágil visto que, ao final da prova o sistema emite um relatório dos resultados e consequentemente reduz o custo, ou seja, no computador o custo para correção e confecção das provas é bem menor do que na prova escrita.

Na prova escrita haverá custo com professores para efetuar a correção e custo com a comissão que irá trabalhar na confecção dos cadernos de prova, além do desperdício em caso de alguma impressão errada.

7.5 Conclusão

Este capítulo apresentou de que forma a prova foi aplicada e como foi feita a análise de dados via TRI. Apresentou ainda, o impacto causado entre os dois modelos de provas, mostrando os pontos positivos e negativos de cada modelo.

8 CONSIDERAÇÕES FINAIS

A inserção da tecnologia já é uma realidade no processo educacional, no entanto ainda precisa ser trabalhada dentro de algumas escolas e melhor aproveitada no processo de avaliação educacional. Apesar de usar o computador na análise de resultados de alguns processos de avaliação, as provas aplicadas ainda são realizadas de forma escrita, o que atrasa o processo de análise dos resultados.

O processo educacional que nos encontramos inseridos nos dias atuais busca estimular no estudante uma visão futurista, que acompanhe a evolução ocorrida em âmbito educacional, que hoje insere a tecnologia de informação e comunicação na evolução da sociedade do conhecimento. O processo educacional sofreu varias mudanças e os alunos devem estar preparados para enfrentar um mundo cheio de desafios e descobertas.

Uma das mudanças sofridas pelo processo educacional foi o uso das Teorias de Medidas no método de avaliação. Anteriormente era muito utilizada a TCM, com o surgimento da TRI alguns sistemas de avaliação passaram a usar a TRI. No entanto as duas ainda são utilizadas.

Assim, neste trabalho de dissertação foi proposto o desenvolvimento de um Módulo de Avaliação Online, que após ser aplicado a um grupo de estudantes, pudesse ter seus resultados avaliados pela TRI, dessa forma por meio da análise dos dados foi feito uma análise do impacto causado pela migração entre a prova escrita e a prova no computador.

O módulo proposto foi aplicado em duas formas, escrito e no computador, com alunos do 5º ano do ensino fundamental, e possibilitou a identificação das principais diferenças causadas pelo impacto dessa migração.

8.1 Contribuições

O trabalho aqui apresentado traz as seguintes contribuições para o universo educacional, mesmo se tratando de um experimento:

- Maiores oportunidades e familiarização dos alunos com as TIC's;
- Automatização do processo de avaliação;

- Desenvolvimento de um Módulo de avaliação aberto para outras áreas de pesquisa;
- Redução de custos na aplicação, visto que não se utiliza prova impressa;
- Agilidade na correção das provas, visto que ao final do teste o sistema emite um relatório com o resultado;
- Interesse maior dos alunos em realizar o teste;
- Interatividade com a prova, visto que no computador as imagens referentes às questões são mais atrativas;
- Calibração e equalização de itens;
- Geração de tabelas com resultado dos parâmetros dos itens.

Este trabalho teve importante contribuição em mostrar por meio de análise feita pela TRI que, para se fazer o teste no computador o grupo de alunos a ser avaliado deve ser preparado para este tipo de teste, ou seja, o aluno deve ter convívio com este tipo de ferramenta.

8.2 Perspectivas Futuras

Durante a realização deste trabalho, percebeu-se a necessidade de dar prosseguimento ao trabalho a partir de novas propostas associadas a ele:

- O Laboratório de Sistemas Inteligentes (LSI) possui um módulo de avaliação Virtual - TANEB em desenvolvimento, e que não possui a TRI no seu processo de avaliação, sendo interessante a inserção deste módulo desenvolvido neste trabalho, no módulo de avaliação VT;
- Integrar ao módulo de avaliação desenvolvido, o CAT (Teste Adaptativo Computadorizado), em que os itens são selecionados conforme o respondente avaliado;
- Aplicar novamente o teste, seguindo os mesmos padrões (escrito e no computador) a um grupo de respondentes também do 5º ano, da rede particular de ensino e comparar com o resultado apresentado neste trabalho;

- Aplicar o teste entre grupos com itens parcialmente distintos, e verificar o impacto causado;
- Dar continuidade a este experimento, envolvendo estudantes com as mesmas características, no entanto pertencentes a diferentes estados e/ou municípios.

REFERENCIAS

1. ANDRADE, D. F., TAVARES, H. R. ;VALLE, R. C. **Teoria da Resposta ao Item**: conceitos e aplicações. SINAPE: SP, 2000.
2. ANDRADE, D. **Instalação**. Disponível em: <http://www.inf.ufsc.br/~dandrade/>. Acesso em: 12 dez. 2012.
3. ANJOS, A.; ANDRADE, D. F. **Teoria de Resposta ao Item com uso do R**. 20° SINAPE, PB – 2012.
4. ANJOS, A. **Estatística básica com uso do Software R**. UFPR , 2010
5. AYDIN, S. **O efeito dos computadores no Teste e Confiabilidade de Inter-Avaliadores de Testes Escritos de Aprendizes de ESL**. TOJET, 2006.
6. BRASIL. Ministério da Educação.**Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Disponível em: <http://portal.inep.gov.br/web/enem/enem>. Acesso em: maio 2012.
7. _____. Ministério da Educação. **ENEM** – Exame Nacional do Ensino Médio . Disponível em: <http://portal.inep.gov.br/web/enem/enem>. Acesso em: 8 maio 2012.
8. CARVALHO, R. M., COELHO, E. G. **Sistema de Testagem Adaptativa Computadorizada (CAT da Universidade de Brasília com Pré-Testagem Automatizada**. Brasília, 2011.
9. ENEM. Disponível em: <http://www.sitedoenem.org/>. Acesso em: 12 set. 2012.
10. FRANCISCO, R. **Aplicação da Teoria de Resposta ao Item (TRI) no Exame Nacional de Cursos (E.N.C) da UNICENTRO**. Curitiba, 2005.
11. FARIAS, A, C. M. O., et al. **Uma Análise da Evolução no Processo de Avaliação Educacional até o uso de Sistemas Inteligentes**. Aveiro, Portugal, 2011.
12. FARIAS, A, C. M. O., TEIXEIRA, M. P. S. S., OLIVEIRA, R. N. B. **Proposta de Desenvolvimento de um Ambiente de Aprendizagem Cooperativa com o uso das Tecnologias de Informação e Comunicação em Educação à Distância para o curso de Sistemas de Informação do Uniceuma**. II SENEPT – Belo Horizonte, 2010.
13. MACHADO, N.J. **Epistemologia e didática**: as concepções de conhecimento e inteligência e a prática docente. 2 ed. São Paulo: Cortez, 1996.

14. PARÂMETROS CURRICULARES NACIONAIS (PCN). **Ensino Médio: Linguagens, Códigos e suas Tecnologias**. Brasília: MEC/SEMT, 1999.
15. PINTO, M. **As novas tecnologias e a educação**. Disponível em: http://www.portalanpedsul.com.br/admin/uploads/2004/Poster/Poster/04_53_48_AS_NOVAS_TECNOLOGIAS_E_A_EDUCACAO.pdf. Acesso em: 20 dez. 2012.
16. SANTOS, M. A. C., SOARES, J. F. **Estimativas não paramétricas da Curvas Características dos Itens em Testes Educacionais**. São Paulo, 2012. Disponível em: <http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1732/1732.pdf>. Acesso em: 25 out. 2012.
17. SGUISSARDI, V. Para avaliar propostas de avaliação do Ensino Superior. In: SGUISSARDI, V. et al. **Avaliação Universitária em questão: reformas do estado e da educação superior**. Campinas, SP: Autores Associados, 1997.cap.5.
18. TEIXEIRA, M. P. S. S. **Un modelo pedagógico para educación a distancia com el uso de las tecnologias de La información y las comunicacione em el Centro Federal de Educación Tecnológica de Maranhão – CEFET/MA**. Tesis apresentada en opción al Grado Científico de Doctor em Ciências Pedagógicas. ICCP, Ciudad da La Habana. 2005.
19. TEZZA, R.; BORNIA, A. C.; **Teoria da Resposta ao Item: Vantagens e Oportunidades para a Engenharia de Produção**. Salvador: ENEGEP 2009.
20. TEZZA, R. **Proposta de um construto para medir usabilidade em sites de e-commerce utilizando a Teoria da Resposta ao Item**. Dissertação de Mestrado. Universidade de Santa Catarina/UFSC. Florianópolis, SC. 2009.
21. VERAS, J. M. **Modelagem para o Software Virtual** - Taneb baseado na Teoria da Resposta ao Item para avaliar o rendimento dos alunos. São Luís, 2010.
22. VENDRAMINI, C. M M.; SILVA, M. C.; CANALE, M. **Análise de Itens de uma prova de Raciocínio Estatístico**. Maringá, 2004.

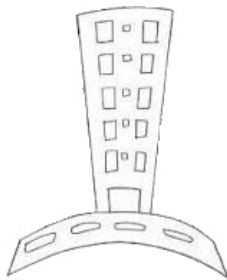
ANEXO

Anexo: Prova escrita aplicada aos alunos do 5º ano fundamental

NOME: _____

ESCOLA: _____

1. Um prédio de 5 andares, possui 10 degraus por andar, para subir até o 4º andar do prédio, Valéria precisará subir quantos degraus?



- a) 20 b) 30 c) 40 d) 50

2. A figura abaixo representa um sólido formado por cubos.



Quantos cubos foram usados para formar o sólido?

- a) 24 b) 27 c) 30 d) 33

3. Vivi tinha 8 lápis no estojo, sendo: 2 vermelhos, 1 preto, 1 azul, 2 amarelos e dois laranjas. Quebraram-se: 1 amarelo, 1 vermelho e 1 preto. Quantos restaram?



- a) 1 vermelho, 1 azul, 1 amarelo e 2 laranjas
b) 1 vermelho, 1 preto, 1 azul e 2 amarelos
c) 2 laranjas, 2 vermelhos e 1 azul
d) 1 amarelo, 1 laranja, 1 preto e 2 vermelhos

4. Observe o desenho da escada abaixo:



O desenho dessa escada oferece formas geométricas. Quais?

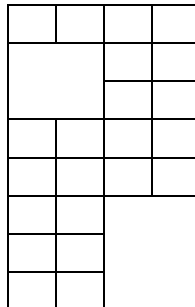
- a) 6 quadrados, 1 retângulo e 1 triângulo
- b) 7 retângulos e 1 triângulo
- c) 5 retângulos, 2 quadrados e 1 triângulo
- d) 9 retângulos e 1 triângulo

5. Das figuras abaixo, quais possuem 5 faces.



- a) Cubo
- b) Paralelepípedo
- c) Pirâmide de base quadrada
- d) Prisma de base hexagonal

6. Carlinhos comprou um caderno quadriculado para estudar matemática, mais ao chegar em casa observou que havia uma página faltando quadrados. Quantos quadrados faltam?



- a) 8
- b) 6
- c) 12
- d) 10

7. A professora de Janaína pediu aos alunos que criassem um desenho com formas geométricas. Quais as formas geométricas utilizadas por Janaína?



- a) Triângulo, Quadrado, Prisma, Círculo
- b) Retângulo, Trapézio, Quadrado, Triângulo

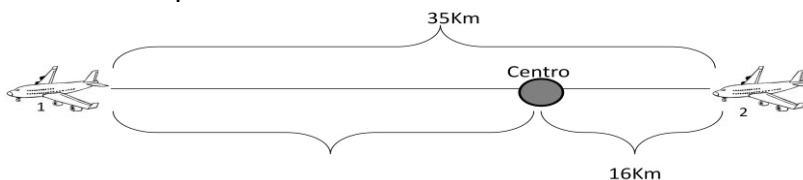
- c) Triângulo, Quadrado, Retângulo, Círculo
 d) Retângulo, Quadrado, Losango, Triângulo

8. Tatiana foi uma em loja e comprou uma caixa de bombons, ao abrir a caixa verificou que os bombons tinham o formato de várias formas geométricas. Quantas formas geométricas há na caixa de bombons?



- a) 3 b) 4 c) 5 d) 6

9. Em uma cidade existem 2 aeroportos distantes um do outro. Observe a figura abaixo e responda:



Qual é a distância entre o aeroporto 1 e o Centro?

- a) 16 km
 b) 19 km
 c) 20 km
 d) 21 km

10. Vanessa precisa saber o peso do seu cachorro. Como ele não fica parado, ela resolveu subir na balança com ele. Vanessa pesa 27kg, observe o peso na balança e responda: Quanto pesa o cachorro de Vanessa?



- a) 8kg
 b) 7kg
 c) 6kg
 d) 5kg

11. Qual desses quadriláteros possuem os 4 ângulos retos?



1



2



3



4



5



6

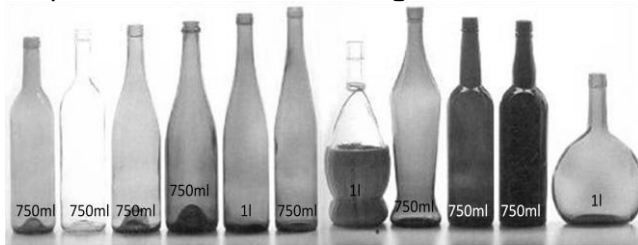


7

- a) 1 e 7
 b) 3 e 6

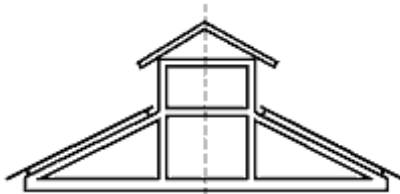
- c) 2 e 4
d) 2 e 5

12. Carlota possui 11 garrafas de diferentes formas e diferentes capacidades. Ela comprou em galão de água de 20 litros, se Carlota encher todas as garrafas, quantos litros restará no galão?



- a) 9L
b) 11L
c) 10L
d) 12L

13. Observe o telhado desta casa.



Quais as formas geométricas eu posso tirar deste telhado?

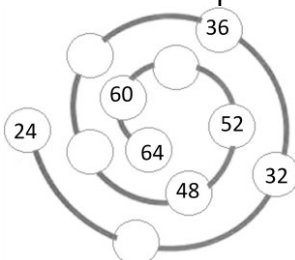
- a) Triângulo, Quadrado e Retângulo
b) Quadrado, Retângulo e Pirâmide
c) Cubo, Triângulo e Retângulo
d) Prisma, Cubo e Pirâmide

14. A Mãe de Marina deu duas moedas, conforme a figura, para ela comprar um pastel com refrigerante. O Copo de refrigerante custa 0,50 e o pastel custa 0,75. Quanto restou do dinheiro de Marina?



- a) 0,15 b) 0,10 c) 0,50 d) 0,25

15. Observe a sequência e responda quais os números faltam ser inseridos.



- a) 28, 40, 44, 56
b) 27, 39, 44, 55

- c) 29, 41, 45, 56
- d) 30, 34, 40, 44

16. Carmem comprou um perfume e ao abrir a embalagem observou que ele tinha o formato de:



- a) Pirâmide
- b) Cone
- c) Prisma
- d) Cilindro

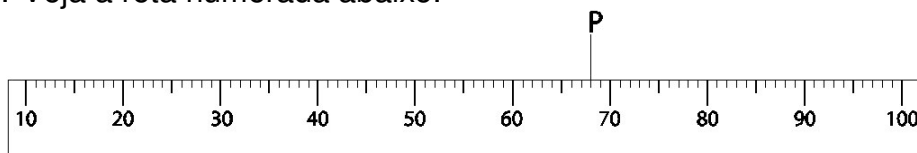
17. Quando Aline colocou um bolo pra assar, o relógio marcava:



O bolo ficou pronto em 40 minutos. Que horário o relógio estava marcando quando o bolo ficou pronto?

- a) 07:20
- b) 07:40
- c) 08:00
- d) 08:10

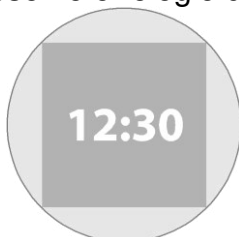
18. Veja a reta numerada abaixo.



Qual é o número representado pela letra P?

- a) 68
- b) 88
- c) 31
- d) 18

19. Observe o relógio digital abaixo.



Que horário aparecerá no mostrador desse relógio daqui a 10 minutos?

- a) 12:20
- b) 12:30
- c) 12:35
- d) 12:40

20. Os amigos Pedro, Renato e Roberto comemoraram a vitória do seu time em uma

pizzaria. Eles dividiram uma pizza em 8 pedaços iguais. Pedro comeu $\frac{3}{8}$, Renato

comeu $\frac{1}{8}$ e Roberto comeu $\frac{2}{8}$ dessa pizza.

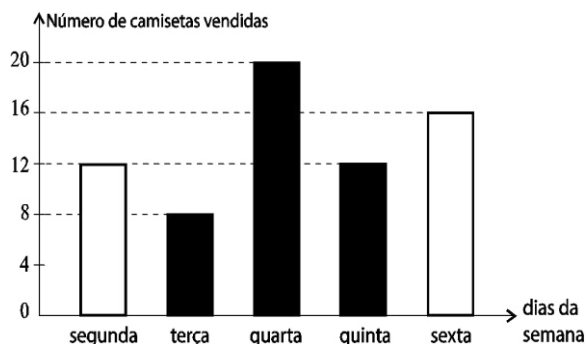
Que fração da pizza eles comeram?

- a) $\frac{6}{8}$
- b) $\frac{4}{8}$
- c) $\frac{6}{14}$
- d) $\frac{6}{24}$

21. Carlos escreveu em seu caderno o número 83 246. Nesse número, o valor posicional do algarismo 3 é:

- a) 3
- b) 30
- c) 300
- d) 3000

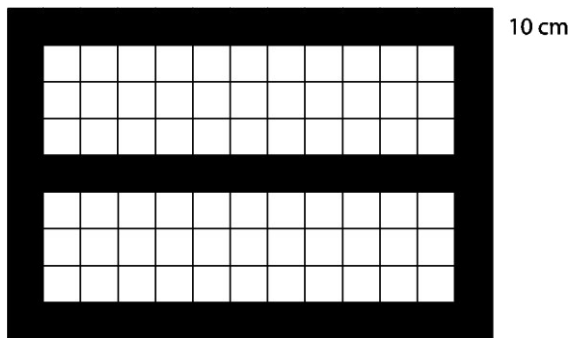
22. O senhor Luiz tem uma loja que vende produtos esportivos. Na semana passada, ele pretendia vender as 100 camisetas de times de futebol que possuía. O gráfico mostra o número de camisetas que foram vendidas na loja do senhor Luiz em cada dia da semana passada, de segunda-feira a sexta-feira.



Após a venda de camisetas nesses cinco dias, quantas sobraram?

- a) 32
- b) 42
- c) 64
- d) 80

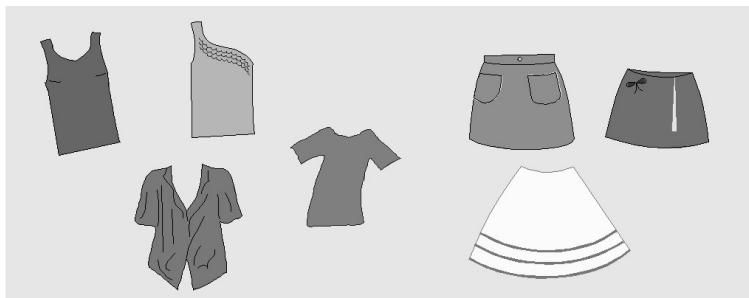
23. Alice montou, com retalhos de tecido, um tapete como o representado abaixo. Cada quadradinho é um retalho com 10 cm de lado.



A área ocupada pelo tapete é:

- a) 2 200 cm²
- b) 4 400 cm².
- c) 8 800 cm².
- d) 11 700 cm².

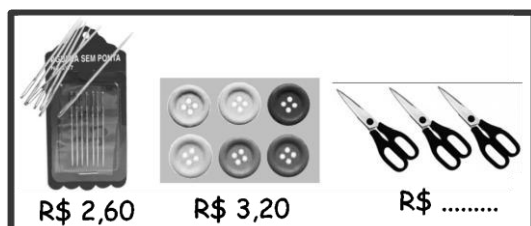
24. Laura separou as roupas abaixo para escolher a que usará na festa de aniversário de sua amiga Bebel.



Escolhendo uma blusa e uma saia, Laura poderá se vestir de

- a) 3 maneiras diferentes.
- b) 4 maneiras diferentes.
- c) 7 maneiras diferentes.
- d) 12 maneiras diferentes.

25. A despesa de Quitéria no Armazinho *Pague e Leve* foi de R\$ 18,70. Ela comprou uma cartela de agulhas, uma cartela de botões e três tesouras.



O preço de cada tesoura é

- a) R\$ 4,30.
- b) R\$ 5,80.

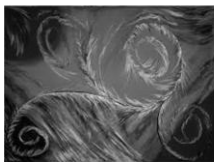
- c) R\$ 6,23.
- d) R\$ 8,60.

26. Um rolo de fio tem 9,9 quilogramas, se um 1 metro de fio equivale a 0,55 quilogramas. Quantos metros tem o rolo?



- a) 16 m
- b) 18 m
- c) 20 m
- d) 22 m

27. Valéria comprou um quadro por R\$ 250,00 e gastou R\$ 50,00 com moldura e R\$ 30,00 com o vidro anti-reflexo. Ela resolveu vender o quadro, mais pra que ela obtenha 20% de lucro em cima do valor gasto, quanto ela deve cobrar pelo quadro?



RS 250,00

+



RS 50,00

+



RS 30,00

- a) R\$ 330,00
- b) R\$ 356,00
- c) R\$ 370,00
- d) R\$ 396,00

28. O nº de telefone é composto por oito algarismos. Considere os algarismos que compõem o número de telefone: 3249-8657. Qual a opção que indica o maior número que estes algarismos podem formar:



- a) 98675423
- b) 98764523
- c) 98765432
- d) 98675432

29. Yasmin está contando as horas para chegar o seu aniversário, se hoje é 08/09 e se ela sabe que faltam 63 dias, qual é o dia do aniversário de Yasmin?



- a) 10/11
- b) 10/10
- c) 08/11
- d) 08/10

30. Um fazendeiro tinha 356 gados, em um leilão ele colocou 20 gados à venda, mais apenas 9 foram arrematados. Se ele arrematou 7, como ficou a situação dele?



- a) Ficou com dois gados a mais
- b) Ficou com dois gados a menos
- c) Ficou com a mesma quantidade de gados
- d) Ficou com três gados a mais

