

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

Luís Carlos Costa Fonseca¹
Lidinalva de Almada Coutinho¹
Lanyllo Araujo dos Santos¹

¹ Programa de Pós-Graduação em Engenharia da Computação e Sistemas –
Universidade Estadual do Maranhão – UEMA

lccfonseca@gmail.com, lidi_almada@hotmail.com, nylg12@hotmail.com

Resumo: Neste artigo, apresentam-se os resultados parciais de uma pesquisa no âmbito da avaliação educacional, com o objetivo de propor uma ferramenta computacional baseada na Teoria de Resposta ao Item, que permita verificar a proficiência de indivíduos. O trabalho em epígrafe é um recorte da pesquisa realizada para a dissertação de mestrado, que tem como problemática central compreender –Como avaliar as competências e habilidades dos discentes do curso de Matemática, do CESC/UEMA, utilizando uma ferramenta computacional que identifique as habilidades dos alunos e forneça um feedback como suporte para a aprendizagem, de acordo com as dificuldades encontradas no teste, empregando técnicas da Teoria de Resposta ao Item (TRI)||.

Palavras-Chave: Teoria de Resposta ao Item, Proficiência, Habilidades.

TOOL TO ESTIMATE INDIVIDUAL'S ABILITY BASED ON ITEM RESPONSE THEORY

Abstract: In this article, we present the partial results of a research field of educational assessment, with the aim of proposing a computational tool based on Response Theory to the Item, to verify the proficiency individuals. The epigraph work is part of a research conducted for the dissertation, which has as its central issue to understand "How to assess the skills and abilities of students of mathematics courses, the CESC / UEMA using a computational tool that identifies the skills of students and provide a feedback as a support for learning, according to the difficulties encountered in testing, employing techniques of the Item Response Theory (IRT) ".

Keywords: Item Response Theory, Proficiency, Skills.

Introdução

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

É indiscutível a necessidade de que sejam reavaliados e aperfeiçoados os métodos de avaliação em todos os segmentos de ensino, uma inquietação evidente nos inúmeros estudos voltados para esta temática, Franco (2001) e Gatti (1996). Desta

forma, o objetivo desta pesquisa é propor uma ferramenta computacional baseada na TRI, que permita verificar a proficiência de indivíduos.

Para tanto, serão levados em consideração alunos do ensino superior, seguindo os moldes exigidos pelo ENADE. Também serão apresentados os fundamentos básicos da TRI, suas vantagens e oportunidades dentro do processo de avaliação educacional; propondo uma modelagem de software que trabalhe baseado nos conceitos da TRI na avaliação da aprendizagem; desenvolvendo uma ferramenta de avaliação de aprendizagem que utilize a TRI e que abstraia a complexidade de sua utilização.

O Curso de Matemática Licenciatura do Centro de Estudos Superiores de Caxias da Universidade Estadual do Maranhão – CESC/UEMA vem passando por grandes dificuldades tanto no âmbito do quadro de docentes efetivos (não atende a demanda), quanto em questões estruturais relacionadas a laboratórios. Refletir e discutir sobre a formação de professores de Matemática tem se tornado cada vez mais frequente, visto que nas três últimas avaliações do ENADE (2005, 2008, 2011), de acordo com relatórios do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP, o Curso tem apresentando resultados insatisfatório, sendo motivo de grande preocupação para toda a comunidade acadêmica do referido Centro.

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

Considerando o exposto, este trabalho se justifica pela necessidade da criação de uma ferramenta computacional de avaliação, utilizando recursos da TRI que contemple um modelo pedagógico que atenda aos anseios da comunidade estudantil do ensino superior, no sentido de familiarizá-los com processo de avaliação do ENADE.

AVALIAÇÃO EDUCACIONAL

Nas últimas décadas, o processo de avaliação educacional vem se destacando, através de iniciativas que permitem a implementação de ações mais condizentes com a oferta de uma educação de qualidade e promoção da equidade de oportunidades educacionais. Nesse sentido, de acordo com Dias Sobrinho (2004), a avaliação é assumida como um elemento que traz implícito o compromisso social, com a garantia de padrões de excelência na formação e na produção de conhecimentos, garantindo a todos que possam se beneficiar do ensino superior; com o desenvolvimento de consciência crítica e com a melhoria das condições de vida da sociedade.

No contexto brasileiro, Castro (2009) afirma que tais tendências têm influenciado na criação do Exame Nacional de Desempenho dos Estudantes (ENADE), Sistema Nacional de Avaliação da Educação Básica (Saeb), no Exame Nacional do Ensino Médio (Enem), na Prova Brasil, além da participação no Programa Internacional de Avaliação de Estudantes (Pisa).

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

Avaliação como uma componente do processo de ensino que visa, através da verificação e qualificação dos resultados obtidos, a

determinar a correspondência destes com os objetivos propostos e, daí, orientar a tomada de decisões em relação às atividades didáticas seguintes. (Libâneo 1991, p.196).

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

Criado em 2004, o ENADE substituiu o Exame Nacional de Cursos (também conhecido como Provão), o exame é realizado obedecendo a um ciclo avaliativo, com duração de três anos, entre os cursos avaliados. É importante enfatizar que num processo avaliativo democrático, os sujeitos envolvidos devem ter consciência sobre qual é o ideal, e entender a serviço de quem está esse ideal. Avaliar é um processo dinâmico de reflexão sobre o que se faz. É um movimento constante e permanente entre ação e reflexão. Desse modo, avaliação e aprendizagem estão intrinsecamente entrelaçados. Uma boa avaliação deve começar, a partir do diagnóstico da realidade. A avaliação qualitativa ultrapassa a avaliação quantitativa, porém uma não pode dispensar a outra.

Entende que no espaço educativo os processos são mais relevantes que os produtos, não fazendo jus à realidade, se reduzida apenas às manifestações empiricamente mensuráveis. (...) A avaliação qualitativa gostaria de chegar até à face qualitativa da realidade, ou pelo menos de se aproximar dela. (Demo 2004, p.156).

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

De acordo com Allen & Yen (2002), há uma preocupação em conceber, desenvolver e aplicar métodos avaliativos que representem de forma mais fiel possível os resultados da aprendizagem. Nesse contexto, a teoria da medida tem tido implicações diretas no delineamento, interpretação e resultados de pesquisas e avaliações educacionais. Para Pasquali (2003), Tal teoria desenvolve uma discussão epistemológica em torno da utilização dos números no estudo científico dos fenômenos naturais.

Produzir um instrumento para medir uma variável nas ciências sociais não é uma tarefa fácil e inclui uma série de etapas que deve ser seguida rigorosamente. Segundo Andrade et al. (2010), deve-se considerar as seguintes etapas: conceituação dos comportamentos que definem operacionalmente o construto em questão, elaboração de itens que acessem o construto, administração dos itens elaborados para amostras pré-definidas, refinamento do instrumento baseado em análises dos itens e realização de estudos de validade e confiabilidade. É importante considerar essas etapas para

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

TEORIA DE RESPOSTA AO ITEM – TRI

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

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

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

Em 1970, Bock & Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas. Na primeira etapa, estimam-se os parâmetros dos itens, assumindo-se certa distribuição para as habilidades. Na segunda etapa, assumindo os parâmetros dos itens conhecidos, estimam-se as habilidades.

Recentemente, métodos bayesianos Moreira Junior (2011), foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens respondidos corretamente ou incorretamente por todos os respondentes, e também o problema da estimação das habilidades dos respondentes que acertaram ou erraram todos os itens da prova.

A Teoria de Resposta ao Item no Brasil

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

A TRI trata o problema da estimação da habilidade e conhecimento de um examinando de forma essencialmente diferente: o enfoque das análises desvincula-se das provas (TCT) e concentra-se nos Itens; se na TCT as estatísticas dos itens dependem da população dos examinandos e da prova a qual os itens pertencem, na TRI cria-se o

conceito de que os parâmetros dos itens, obtidos no processo estatístico de "calibração" dos parâmetros de dificuldade, discriminação e acerto casual dos itens Valle (1999), são características próprias dos mesmos.

Modelos da Resposta ao Item

De acordo com Andrade et al. (2000), entre os modelos propostos na literatura, consideram-se fundamentalmente três fatores: a natureza do item, o número de populações envolvidas e a quantidade de traços latentes que está sendo medida.

Na educação, os modelos de TRI mais utilizados são os modelos logísticos para itens dicotômicos, que podem ser utilizados tanto para a análise de itens de múltipla escolha dicotomizados, quanto para a análise de itens abertos, quando avaliados também de forma dicotomizada e diferenciam-se pelo número de parâmetros utilizados para descrever o item: 1, 2 e 3 parâmetros, que consideram, respectivamente, a dificuldade do item, a dificuldade e a discriminação, e a dificuldade, discriminação e probabilidade de acerto casual ou -chute||.

Dos modelos propostos pela TRI, o modelo logístico unidimensional de 3 parâmetros (ML3) é atualmente o mais utilizado e é dado por:

$$P_{ij} = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}} \quad (1)$$

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

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

INFERÊNCIA BAYESIANA

A inferência estatística trata dos procedimentos para fazer generalizações sobre as características de uma população com base nas informações contida na amostra. Para tanto, tem-se a abordagem clássica e a bayesiana. Na abordagem clássica, as inferências são feitas a partir de informações fornecidas por uma amostra aleatória retirada da população enquanto que na abordagem bayesiana deve ser considerada a incorporação de informações obtidas de experimentos passados ou experiências de especialistas. Para Pollard (1986), a aproximação bayesiana pode ser compreendida como uma extensão da aproximação clássica.

Na inferência bayesiana, qualquer informação sobre o parâmetro deve ser considerada e incorporada aos dados, através da distribuição a priori. A distribuição dos dados é chamada de função de verossimilhança. O'Hagan (1994), citado por Bragion et al. (2006), a distribuição resultante entre a junção da distribuição a priori com a função de verossimilhança, é chamada distribuição a posteriori e é com base nela que é feita toda inferência bayesiana.

Redes Bayesianas

A aplicação da inferência bayesiana sobre um número pequeno de variáveis relacionadas é um processo relativamente simples. Mas, em situações reais em que um grande número de variáveis e estados é encontrado a inferência pode não ser trivial.

Em redes Bayesianas a representação das variáveis e relações é feita utilizando Teoria dos Grafos. As variáveis são os nós e os arcos identificam as relações entre as variáveis, formando um grafo dirigido e sem ciclos.

MATERIAL E MÉTODO

Nesta seção, delinea-se o aporte teórico metodológico utilizado para implementação desta pesquisa.

Arquitetura geral de um Sistema de Provas TRI

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

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

A figura abaixo ilustra a Arquitetura do Sistema, ela é composta de três módulos. O primeiro é a ferramenta de autoria de provas, acessada pelo ator autor, com seus casos de uso ilustrados na Figura 2. O segundo é a ferramenta de aplicação de provas, que é onde o respondente terá acesso às provas do sistema. O último é o corretor de provas, uma ferramenta computacional que implementa modelo bayesiano de correção de provas.

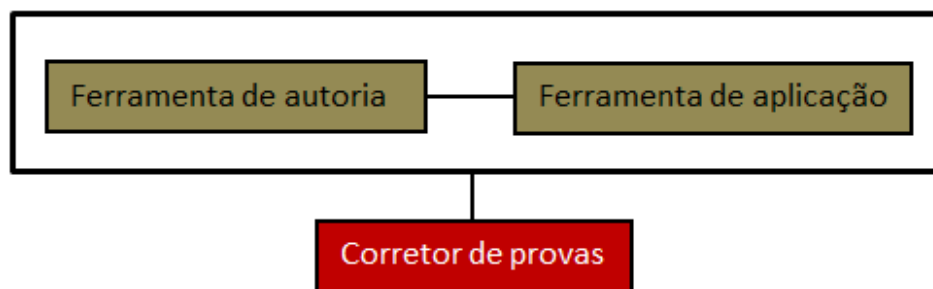


Figura 1: Arquitetura Geral do Sistema

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

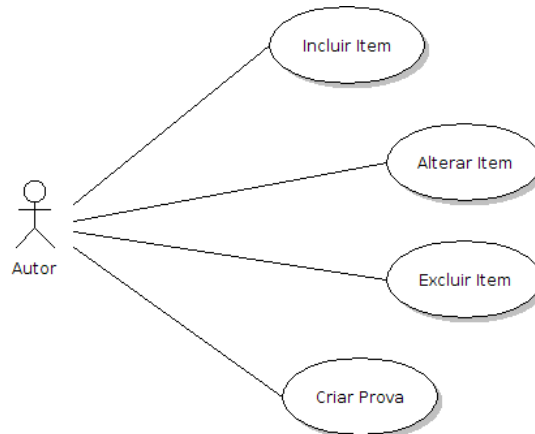


Figura 2: Caso de Uso Autor

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

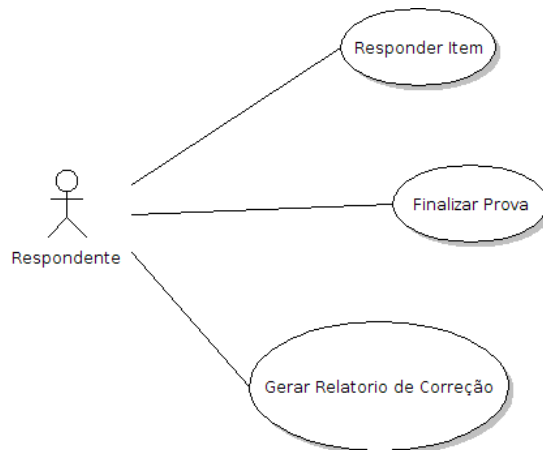


Figura 3 Caso de Uso Aluno

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

Uma proposta de corretor de provas TRI

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

Para entender melhor o problema, ter-se-á o seguinte exemplo:

Uma escala de habilidades busca interpretar qualitativamente os valores obtidos pelos cálculos realizados por fórmulas da TRI, isto é, a escala é construída com a finalidade de saber quais os conceitos de uma determinada área, que os alunos sabem e quais eles ainda não aprenderam (Valle, 1999). Por exemplo, na escala de habilidades do SAEB para conceitos de matemática do 5º e 9º anos, um aluno que acertar questões de nível 7, presume-se que ele aprendeu ou já sabe os seguintes conceitos:

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

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

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

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

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

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

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

1. $P(U = 1 | = Z) = 0,65$ e
2. $P(U = 1 | = Y) = 0,50$ e
3. $P(U = 1 | = Z) - P(U = 1 | = Y) = 0,30$

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

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

CONSIDERAÇÕES FINAIS

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

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

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

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

REFERENCIAS BIBLIOGRÁFICAS

ALLEN, M. J. & Yen, W. M. **Introduction to measurement theory**. Illinois: Waveland Press. 2002;

ANDRADE, Dalton Francisco, TAVARES, Heliton Ribeiro & VALLE, Raquel da Cunha. **Teoria da Resposta ao Item**: Conceitos e Aplicações. SINAPE. 2000;

ANDRADE, Josemberg Moura de; LAROS, Jacob Arie; GOUVEIA, Valdeiney Veloso. **O uso da teoria de resposta ao item em avaliações educacionais:** diretrizes para pesquisadores. Aval. psicol., Porto Alegre, v. 9, n. 3, dez. 2010. Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000300009&lng=pt&nrm=iso>. acesso em 11 mar. 2013;

BRAGION, Maria de Lourdes Lima. **Um modelo de Teoria de Resposta ao Item para os dados do vestibular 2006-2 da UFLA.** Disponível em: http://www.mch.ifsuldeminas.edu.br/~biblioteca/biblioteca_digital/Documentos/Doc_Maria%20de%20Lourdes%202.pdf. Acesso em: 05/03/2013;

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

CASTRO, M.H.G. **Sistemas de avaliação da educação no Brasil:** avanços e novos desafios. São Paulo em Perspectiva, São Paulo, Fundação Seade, v. 23, n. 1, p. 5-18, jan./jun. 2009. Disponível em: < http://www.seade.gov.br/produtos/spp/v23n01/v23n01_01.pdf. Acesso em: 06/05/2013;

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

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

ENADE 2005. Disponível em: <http://enadeies.inep.gov.br/enadeIes/enadeResultado/>. Acesso: 15/04/2013;

ENADE 2011. Disponível em: <http://enadeies.inep.gov.br/enadeIes/enadeResultado/>. Acesso: 15/04/2013;

ENADE 2008. Disponível em: <http://enadeies.inep.gov.br/enadeIes/enadeResultado/>. Acesso: 15/04/2013;

FRANCO, Creso. O SAEB — sistema de avaliação da educação básica: potencialidades, problemas e desafios. Revista Brasileira de Educação, 17, pp. 127 - 133, 2001;

GATTI, Bernardete A. **Avaliação de sistemas educacionais no Brasil.** Disponível em: <http://sisifo.fpce.ul.pt/pdfs/Revista%209%20PT%20d1.pdf>. Acesso em: 06/05/2013;

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

LIBÂNIO, José Carlos. **Didática.** São Paulo, Editora Cortez, 1991;

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

MARQUES, R. L.; DUTRA, I. **Redes Bayesianas:** o que são, para que servem, algoritmos e exemplos de aplicações. Rio de Janeiro: [s.n.], 2008. Disponível em: www.cos.ufrj.br/~ines/courses/cos740/leila/cos740/Bayesianas.pdf. Acesso em: 08/04/2013;

MOREIRA JUNIOR, Fernando de Jesus. **Sistemática para implantação de testes adaptativos informatizados baseados na teoria da resposta ao item.** Disponível em:

<http://repositorio.ufsc.br/bitstream/handle/123456789/95506/299657.pdf?sequence=1>. Acesso em: 08/04/2013;

NEAPOLITAN, R. E. **Learning Bayesian Networks**. [S.l.]: Prentice Hall, 2003;

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

PEREIRA.C.A.;LABIDIS; FONSECA.L.C.C. **Modelagem do Sistema de Avaliação de Conhecimento Aplicável aos Cursos de Graduação, Segundo Parâmetros do ENADE**. Disponível em <http://periodicos.unesc.net/index.php/sulcomp/article/view/286/294> acesso em: 08/04/2013;

PEREIRA, Cleber Augusto. **Modelagem do sistema de avaliação de conhecimento, segundo parâmetros do enade, aplicável aos cursos superiores de Graduação: uma proposta quanto a forma de avaliação nas IES**. Disponível em: http://www.tedebc.ufma.br/tde_busca/arquivo.php?codArquivo=516. Acesso em: 20/03/2013;

POLLARD, W. E. **Bayesian statistics for evaluation research An Introduction**. Sage Publications New Delhi. 241p. 1986;

RAMOS, Edla Faust. **O Papel da Avaliação Educacional nos Processos de Aprendizagem Autônomos e Cooperativos**. In LISIGEN, Irlan, et. al. Formação do Engenheiro: Desafios da atuação docente, Tendências Curriculares e Questões da Educação Tecnológica . Florianópolis: Editora da UFSC, 1999, p. 207 a 228;

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

VERAS, Jaclason Machado. **Modelagem para o software Virtual-TANEB baseado na Teoria da Resposta ao item para avaliar o rendimento dos alunos**. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetaileObraForm.do?select_action=&co_obra=169345. Acesso em 22/03/2013.