

## TESTES ADAPTATIVOS INFORMATIZADOS BASEADOS EM TEORIA DE RESPOSTA AO ITEM UTILIZADOS EM AMBIENTES VIRTUAIS DE APRENDIZAGEM\*

Fabrcia Damando Santos\*\*

Leonardo Guerra de Rezende Guedes\*\*\*

**Resumo:** A utilizaço de ambientes virtuais de aprendizagem para educaço a distncia tem se expandido nos ltimos anos, assim como a utilizaço de softwares para esses novos ambientes educacionais. Neste artigo iremos apresentar uma ferramenta computacional para realizaço de avaliaço adaptativa em cursos a distncia utilizando como fundamentaço a teoria de resposta ao item e os testes adaptativos informatizados. Apresentaremos os resultados obtidos e as concluses advindas com este trabalho.

**Palavras Chaves:** *Educaço a Distncia, Avaliaço, Teoria de Resposta ao item, Testes Adaptativos Informatizados.*

**Abstract:** The utilisation of virtual environment of learning to distance education have been expended in the last years. In this article it will be presented a software tool to the adaptative avaliation in web course based on the item answer theory and the computerized adaptative tests. Will present the obtained results and the conclusions given by this work.

**Key-words:** distance education, avaliation, item response theory, computerized adaptative tests.

### 1. Introduço

A avaliaço tem sido palco de grandes discusses e estudos no ensino, tanto no presencial quanto no a distncia. Novas prticas avaliativas contestam a avaliaço como significado social e poltico. Evitar que a avaliaço julgue pela comparaço e classificaço  uma busca constante. Para os cursos a distncia essa preocupaço tambm existe e, para isso o desenvolvimento de softwares aplicados na avaliaço de alunos, que forneçam informaçes de acompanhamento do aluno para o professor tem a finalidade de aprimorar o ensino atravs da Internet.

A proposta apresentada  uma ferramenta computacional destinada ao professor, que receber informaçes sobre as habilidades dos alunos, os nveis de dificuldade das questes, acompanhamento do aluno no teste e as notas ponderadas equalizadas. Esta ferramenta  implementada a partir da Teoria de Resposta ao item e das metodologias dos Testes Adaptativos Informatizados, pela detecço e classificaço de dificuldades e habilidades dos alunos e que serviro principalmente como instrumento de auto-avaliaço para o professor.

### 2. Teoria de Resposta ao Item

---

\* Artigo extrado da dissertaço de Mestrado: Ferramenta Avaliativa para apoio Pedaggico Baseada em Testes Adaptativos Informatizados e Teoria de Resposta ao Item. UFG. Ano de obtenço: 2003

\*\* Mestre em Engenharia da Computao. Escola Tcnica da Universidade Federal do Rio Grande do Sul. Professora Substituta – fdamando@yahoo.com.br.

\*\*\* Doutor em Engenharia Eltrica. Professor Adjunto Universidade Federal de Gois e Universidade Catlica de Gois.leonardo.guedes@uol.com.br – www.leonardoguedes.com.br

A Teoria de Resposta ao Item é uma reunião de modelos estatísticos usados para fazer predições, estimativas ou inferências sobre as habilidades (ou competências) medidas em um teste. Através dos modelos estatísticos é possível prever tais habilidades por meio de correspondências entre a pontuação obtida por um estudante em uma situação de teste e os itens a ele fornecido [Hambleton & Swaminathan, 1985; Rudner, 1998].

A Teoria de Resposta ao Item propõe modelos que representam a relação entre a probabilidade de uma resposta certa a um item e a habilidade de um aluno, nos proporcionando avaliar individualmente um aluno, pois, cada estudante responderá itens referente à sua habilidade, tornando a avaliação personalizada.

A Teoria de Resposta ao Item possui vários modelos a serem estudados, dentre os modelos propostos pela teoria, optamos por utilizar o Modelo Logístico de um Parâmetro, também conhecido como “The Rasch”. Que foi desenvolvido pelo matemático dinamarquês Georg Rasch [Baker, 2001]. Cujas equação é dada por:

$$P_i(\theta) = \frac{1}{1 + \exp^{-1(\theta - b_i)}}$$

onde:

1.  $P_i(\theta)$  é a probabilidade de um determinado aluno com habilidade  $\theta$ , responder a um item  $i$  corretamente;
2.  $b_i$  é o índice de dificuldade do item.

O parâmetro  $b_i$  pode ser alterado à medida em que os estudantes passam a realizar o teste e a responder corretamente ou incorretamente ao item. Os valores para  $b_i$  variam neste modelo de -2.0 a +2.0, sendo que valores próximos a -2.0 são considerados itens fáceis e os valores próximos a +2.0 são considerados itens difíceis.

A partir do Modelo estudado desenvolvemos uma nova proposta para o cálculo da pontuação obtida por um aluno e a sua habilidade.

### 3. Proposta de Modelos de Pontuação e Habilidade

Este trabalho propõe uma nova abordagem de pontuação e habilidade do aluno. Nestas abordagens utilizamos os conceitos de Esperança e Mediana.

Para realizar os procedimentos de pontuação de um aluno em um teste, quanto se tornar necessário a apresentação de notas para a instituição, exige-se a responsabilidade de transformar a habilidade do aluno em classificação.

No nosso caso, devemos calcular a Esperança de acerto de uma questão de nível de dificuldade  $b_i$  sendo  $P_i(\theta)$  a probabilidade de acerto da questão por um aluno com nível de habilidade  $\theta$ .

$$E(x) = b_i \cdot P_i(\theta)$$

Para calcular a nota  $\pi$  do aluno, iremos calcular a média das esperanças de todas as questões respondidas corretamente e dividir pelo cálculo da média das esperanças de todas as questões (corretas e incorretas) respondidas. Onde termos a seguinte equação:

$$\pi = \frac{\sum_{i=1}^n R_i \cdot E_i(x)}{\sum_{i=1}^n E_i(x)} = \frac{\sum_{i=1}^n R_i \cdot b_i \cdot P_i(\theta)}{\sum_{i=1}^n b_i \cdot P_i(\theta)}$$

Onde  $R_i$  representa a resposta do item  $i$  (0 ou 1).

Para se estimar a habilidade inicial de um aluno, podemos utilizar um valor “default”, e a partir desse valor ajustar a habilidade do aluno, na qual se obtém através das respostas dos itens do teste.

A habilidade do aluno é obtida a partir da probabilidade de acerto de uma questão de determinado nível de dificuldade  $b_i$ . Um aluno com habilidade  $\theta$  terá 50% de probabilidade de acerto de questões com nível de dificuldade  $b_i = \theta$ .

Assim, devemos encontrar qual nível de dificuldade de questão  $b_i$  o aluno terá 50% de chances de acerto, que será realizada pela busca da mediana das questões aplicadas, ou seja, o nível de habilidade tal que metade das questões aplicadas estarão acima deste nível de metade abaixo.

Desenvolvemos uma nova fórmula para cálculo da habilidade dos alunos. Para a Teoria de Resposta ao Item, a habilidade varia de -2.0 a +2.0. Como no nosso caso o índice de dificuldade de um item pode variar de 0 a 10, ajustamos a fórmula da habilidade para:

$$b_i = (b_{i \cdot} / 2,5) - 2$$

onde:

1.  $b_i$  é o índice de dificuldades dos itens
2.  $b_{i \cdot}$  é o nosso valor do nível de dificuldade (0 a 10)

Para calcularmos a habilidade  $\theta$ , sempre trabalhamos com o valor mediano do itens aplicados ao teste, o que nos leva a um valor central.

$$\theta = [(med_{niveis} / 2,5) - 2]$$

Portanto, aplicando o valor da habilidade no Modelo The Rash, encontramos a probabilidade de um aluno responder corretamente um item dado a sua habilidade.

#### 4. Modelos Desenvolvidos

Para o desenvolvimento dos modelos avaliativos, utilizamos um banco de itens (questões) com vários níveis de dificuldade, variando de 0 a 10, onde as questões próximas a 0 são consideradas fáceis e as próximas a 10 consideradas difíceis. As questões são cadastradas pelo próprio professor, composta por um código, descrição da pergunta, cinco (05) alternativas, o nível de dificuldade inicial da questão e o gabarito (resposta correta).

Todos os alunos iniciam o teste com uma questão de mesmo nível de dificuldade. No nosso caso, iniciamos um teste com uma pergunta de nível de dificuldade 5, que é considerada de nível de dificuldade intermediário. Sendo assim, consideramos também que todos os alunos possuem a mesma habilidade, mas relembramos que esta habilidade irá se ajustar no decorrer do teste.

O *Modelo Avaliativo I* seleciona um item após o outro, ou seja, a escolha do próximo item a ser administrado é realizada logo após um item ser respondido. Administrado o primeiro item, caso o aluno o responda corretamente será fornecido a ele um novo item com um nível de dificuldade maior. Caso o aluno responda incorretamente, o próximo item a ser administrado a ele será um item de nível de dificuldade inferior ao que estava fazendo.

Sejam dois alunos fazendo o teste ao mesmo tempo, ambos provavelmente não responderão as mesmas questões, pois elas serão administradas de acordo com a habilidade de cada um.

O *Modelo Avaliativo II*, também inicia o teste no mesmo nível de dificuldade. Para o Modelo Avaliativo II, dado que um aluno iniciou o teste e respondeu corretamente ao item, o próximo item a ser administrado será calculado pela média do nível do item atual e o nível máximo admitido no teste, como abaixo:

$$n_{\text{novo}} = \frac{n_{\text{atual}} + n_{\text{máx}}}{2}$$

Caso contrário, se o aluno responder incorretamente ao item, o próximo item a ser administrado ao aluno será feito de duas maneiras:

Para o primeiro erro, calcula-se o nível do próximo item fazendo a média entre o nível do item atual e o nível do item anterior respondido pelo aluno, dessa forma obtemos um item que esteja entre os níveis de dificuldade no qual o aluno está apto a realizar.

$$n_{\text{novo}} = \frac{n_{\text{atual}} + n_{\text{anterior}}}{2}$$

Caso o aluno responda incorretamente a dois itens consecutivos, o próximo item a ser administrado será calculado pela média entre o nível do item atual e o nível do item mínimo, exemplificado pela fórmula abaixo:

$$n_{\text{nov}} = \frac{n_{\text{atual}} + n_{\text{mín}}}{2}$$

#### 4.1. Variação dos níveis de dificuldade de uma questão

A medida em que as questões são respondidas, os níveis de dificuldade podem variar, o que deixa os modelos desenvolvidos auto-ajustáveis ao longo do tempo. O nível de dificuldade de uma questão foi cadastrado pelo professor, mas a medida que a questão passa a ser respondida por um aluno ou turma, o seu nível de dificuldade pode ser modificado.

Assumimos que, quanto mais alunos acertarem a uma determinada questão, ela poderá ser considerada mais fácil e conseqüentemente o seu nível de dificuldade deverá diminuir. No mesmo caso, quanto mais alunos responderem incorretamente uma questão, consideramos que esta questão é considerada de nível mais difícil e conseqüentemente o seu nível deverá aumentar.

O novo nível da questão é calculado através da seguinte equação:

$$\text{nível} = \left\{ 10 - \left[ \left( \frac{\text{num\_acertos}}{\text{num\_acessos}} \right) * 10 \right] \right\}$$

Com esses resultados obtemos parâmetros para traçar o perfil de uma turma, ou aluno, auxiliando o professor a diagnosticar os pontos fortes e fracos.

### 5. Ferramenta Computacional

Essa Ferramenta é baseada em Testes Adaptativos Informatizados e Teoria de Resposta ao Item, podendo ser realizada via Web. A Ferramenta oferece a edição de questões, a correção dos testes aplicados aos alunos e posterior divulgação dos resultados. Podendo ser utilizado pelo administrador do sistema, pelo professor e pelo aluno. A Ferramenta foi desenvolvida em linguagem ASP(Active Server Pages) , interface HTML, a validação dos campos através de JavaScript e do banco de dados Access.

O professor é responsável pelo cadastro, alteração e exclusão de questões. O aluno é cadastrado no sistema pelo administrador, recebe um senha e poderá utilizar a ferramenta. Quando o aluno “logar” no sistema para realizar o teste, caberá a ele a escolha do Questionário I (Modelo Avaliativo I) ou o Questionário II (Modelo Avaliativo II). Após a escolha, as questões serão administradas uma a uma estabelecendo os modelos de busca descritos anteriormente.

#### 5.1. Apresentação dos resultados(feedback)

Ao final do teste, um feedback instantâneo é mostrado ao aluno, onde as questões que foram respondidas incorretamente são listadas. Para o professor os resultados se mostram mais consistentes. O professor poderá visualizar o desempenho de cada aluno ou da turma, em forma de nota, listagem das questões respondidas pelos alunos e as “notas” sugeridas. O feedback mostra ainda a mediana, a habilidade do aluno e média

de todas as notas dos alunos no teste. A ferramenta foi desenvolvida para oferecer informações tanto para o professor quanto para o aluno. Simplificar o trabalho de correção de questões para dar apoio ao professor.

A forma de identificar a habilidade do aluno contribui para identificar suas dificuldades e facilidades, fazendo com que dentro de seus limites, qualquer aluno consiga realizar o teste, que será ajustado ao seu perfil.

## 6. Análise dos resultados

Selecionamos alguns alunos e as curvas obtidas dos níveis de dificuldade realizados por cada um, para uma melhor visualização e análise dos dados obtidos.

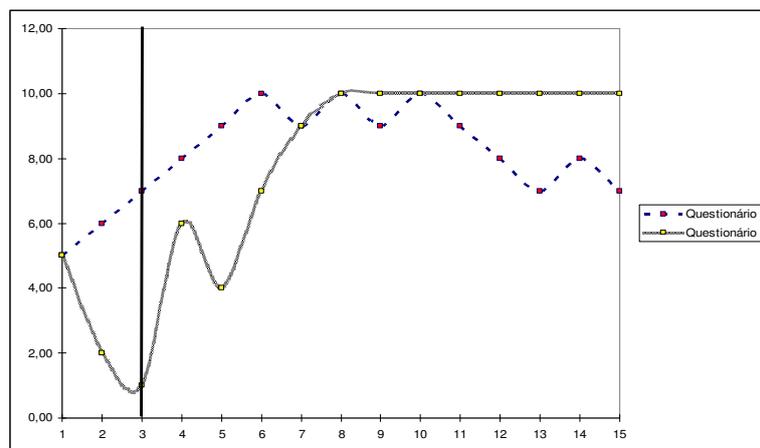


Figura 1- Curvas das habilidades para o Aluno X

Nota-se neste caso que o Modelo Avaliativo I (Questionário I) apresenta aumentos gradativos dos níveis de dificuldade das questões até o aluno alcançar o nível mais alto, somente depois é que encontramos algumas variações alternadas para os níveis, e uma posterior queda gradativa. Para o Modelo Avaliativo II (Questionário II), há uma variabilidade maior nos níveis inicialmente e logo depois há uma estabilização.

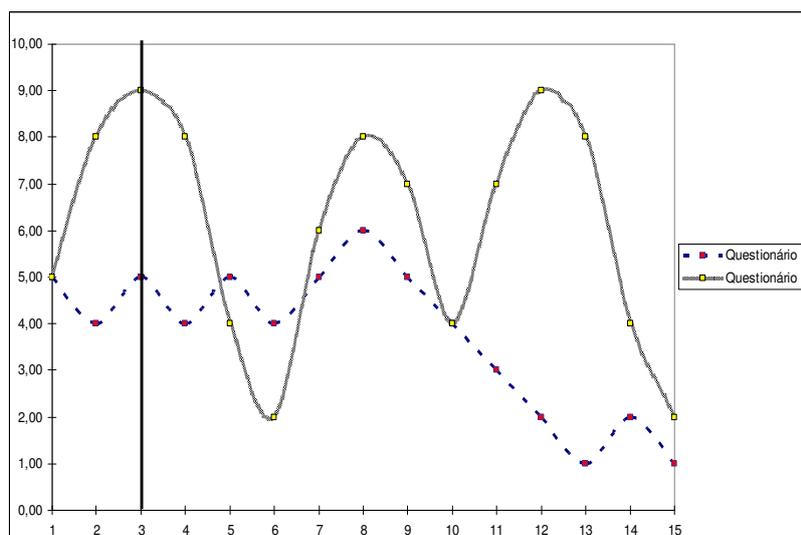


Figura 2- Curvas das habilidades para o Aluno Y

Em todas as curvas, calculamos a habilidade dos alunos, e também calculamos a habilidade eliminando as três primeiras questões respondidas pelos alunos, pois consideramos utilizadas para a familiarização com a ferramenta, para alcançarmos resultados mais precisos.

O Modelo Avaliativo I não submeteu os alunos à uma grande variabilidade de níveis de dificuldade das questões, apresentando-se mais estável e desafiando o aluno em seu nível de habilidade. Já para o modelo avaliativo II, houve uma maior variabilidade de nível de dificuldade, potencializando o acerto.

No Modelo Avaliativo I os níveis de dificuldade variam menos, e pode ser bem aplicado para alunos onde as suas habilidades são intermediárias (nem muito alta e nem muito baixa), pois o nível de uma questão varia gradativamente. Porém, nos casos com alunos de habilidades altas, a mudança de nível é menos rápida, o que leva ao aluno responder questões nas quais ele já possui conhecimento, sendo consideradas fáceis.

Para o modelo avaliativo II, chega-se mais rápido à habilidade do aluno nos casos em que o aluno tem conhecimento avançado ou básico. Se o aluno possui habilidade alta, o modelo II administrará questões de maior nível de dificuldade de forma mais rápida, pois sempre trabalha com a média dos níveis das questões (atual e máxima). Caso o aluno possua uma habilidade baixa, o modelo II também direcionará de forma mais rápida à questões de nível menor.

Quanto aos níveis de dificuldade das questões, eles variaram no decorrer do teste tornando o banco de itens mais calibrado, ou seja, questões com níveis de dificuldade “reais” para um aluno ou turma. Em média 44% das questões aumentaram o nível de dificuldade e 42% diminuíram o seu nível de dificuldade, o que realmente demonstra que estão se ajustando ao longo do tempo.

### **6.1. Cálculo da Correlação**

Realizamos o cálculo da correlação entre as habilidades e notas obtidas pelos alunos nos Questionário I e II. Conseguimos o valor de 0,94 para o Questionário I e de 0,97 para o Questionário II, o que demonstra matematicamente resultados satisfatórios.

## **7. Conclusão**

A Ferramenta desenvolvida pode ser utilizada para acompanhamento do aluno no curso, auxiliando o professor a diagnosticar o desempenho do aluno e se o conteúdo assimilado pelo aluno foi necessário a seu aprendizado.

A capacidade de fornecer subsídios de resposta para a instituição, principalmente para o professor e para o próprio aluno que se submeter ao questionário, fornecendo a retroalimentação, estimulando o aprendizado do aluno, mostrando informações que os auxilie a julgar suas estratégias individuais de estudo e também auxiliando o professor a diagnosticar qual parte do curso precisa de atenção especial, ou seja, coopera no sentido de sensibilizar para o que deve ser mudado ou melhorado.

Quanto aos dois modelos propostos, observou-se que o modelo com menos variações (Modelo I) pode ser bem aplicado a todos os alunos e principalmente àqueles com nível de habilidade média, pois vai testando o aluno nível a nível. Para os casos de

alunos avançados, o Modelo II se apresenta mais favorável, pois faz com que o aluno responda questões com níveis de dificuldade mais compatíveis com a sua habilidade, o que pode favorecê-lo.

## 8. Referência Bibliográfica

[Baker, 2001] Baker, F. B. The Basis of Item Response Theory. Second Edition. ERIC Clearinghouse on Assessment and Evaluation, 2001.

[Hambleton and Swaminathan, 1985] Hambleton, R. . and Swaminathan, H. Response Theory: Principles an Application. Kluwer /nijhoff Publishing, 1985.

[Rudner, 1998] Rudner, M. L. An On-Line, interactive, computer adaptive testing mini-tutorial, 1998. <http://www.ericae.net/scripts/cat/catdemo.html>. Acessado: Mai, 2002.