

TEMA EM DESTAQUE

O EFEITO DA PROVA NA ESTIMATIVA DA PROFICIÊNCIA ATRAVÉS DA TRI

DALTON FRANCISCO DE ANDRADE,
ADRIANO FERRETI BORGATTO

RESUMO

O objetivo deste trabalho foi mostrar a importância da construção da prova para a estimativa da proficiência dos respondentes através da Teoria da Resposta ao Item – TRI, usando uma escala com média 250 e desvio-padrão 50. Foi mostrado que um teste com informação na região da proficiência do respondente fornece estimativas mais precisas do que testes com informações em outras regiões da proficiência. Além disso, mostrou-se também a importância de testes maiores na estimação da proficiência.

PALAVRAS-CHAVE: TESTES, ANÁLISE DE ERROS • SIMULAÇÃO
• TEORIA DE RESPOSTA AO ITEM.

RESUMEN

El objetivo de este trabajo fue mostrar la importancia de la construcción de la prueba para estimar la competencia de quienes responden a través de la Teoría de la Respuesta al Item (TRI), usando una escala con media 250 y desvío patrón 50. Se mostró que un test con informaciones en la región de la competencia de quien responde provee estimativas más precisas que testes con informaciones en otras regiones de la competencia. Además, se mostró la importancia de testes más largos para estimar la competencia.

PALABRAS CLAVE: TESTES • ANÁLISIS DE ERRORES • SIMULACIÓN • TEORÍA DE LA RESPUESTA AL ITEM.

ABSTRACT

The aim of this study was to show the relevance of test building in estimating the proficiency of respondents through the Item Response Theory -IRT, using a scale with a mean of 250 and a standard deviation of 50. It was shown that a test with information in the respondent's proficiency region provides more accurate estimates than tests with information in other areas of proficiency. In addition, the importance of larger tests in estimating proficiency was also shown.

KEYWORDS: TESTING • ERROR ANALYSIS • SIMULATION • ITEM RESPONSE THEORY.

1 INTRODUÇÃO

O Inep vem obtendo informações sobre o desempenho dos alunos do Ensino Fundamental e do Ensino Médio desde 1995, por meio do Sistema Nacional de Avaliação da Educação Básica (Saeb). Este sistema avalia o desempenho dos estudantes em Língua Portuguesa e Matemática, através da aplicação de testes educacionais a cada dois anos.

O desempenho dos estudantes é avaliado conjuntamente entre as séries através da Teoria da Resposta ao Item, usando uma escala única do Saeb com média 250 e desvio-padrão 50, onde a média 250 representa o desempenho médio dos alunos da 8ª série do Ensino Fundamental de 1997.

O objetivo deste trabalho é mostrar a importância da construção da prova para a estimativa da proficiência dos respondentes através da Teoria da Resposta ao Item – TRI, usando uma escala com média 250 e desvio-padrão 50 para tentar reproduzir a situação que pode ocorrer com o Saeb. Em outras palavras, o objetivo é analisar o efeito da prova tomando como referência a curva de informação do teste (prova) e também o tamanho do teste.

2 SIMULAÇÃO

Na simulação obtiveram-se respostas de 1.000 indivíduos representando amostras aleatórias de três populações com distribuição Normal - $N(m,s)$, descritas a seguir:

População 1: $N(200,25)$;

População 2: $N(300,25)$;

População 3: $N(250,50)$.

As duas primeiras populações, tiveram a variabilidade de meio desvio-padrão na escala (250,50), caracterizando assim, populações mais homogêneas do que a população 3. A primeira população teve as proficiências simuladas com um desvio padrão da escala abaixo da média 250, reproduzindo padrões de respostas de alunos com baixa proficiência e a segunda população teve as proficiências simuladas com um desvio padrão da escala acima da média 250, reproduzindo padrões de respostas de alunos com alta proficiência.

Cada uma das populações teve respostas simuladas em seis provas com graus de dificuldade diferentes – de acordo com o parâmetro b do modelo de Rasch (modelo logístico de um parâmetro) – e número de itens diferentes – 24 e 36 itens, descritas a seguir:

Prova 1: Parâmetro b entre 150 e 400 com 24 itens;

Prova 2: Parâmetro b entre 250 e 400 com 24 itens;

Prova 3: Parâmetro b entre 150 e 300 com 24 itens;

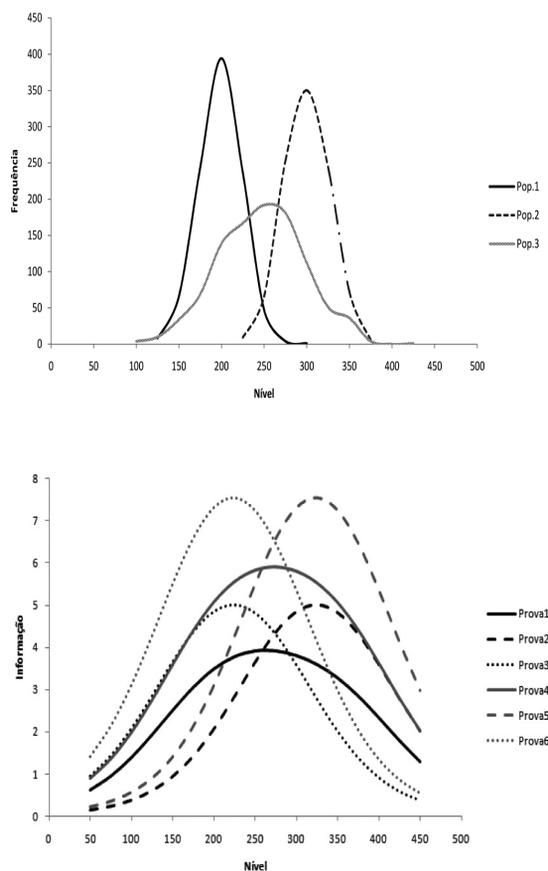
Prova 4: Parâmetro b entre 150 e 400 com 36 itens;

Prova 5: Parâmetro b entre 250 e 400 com 36 itens;

Prova 6: Parâmetro b entre 150 e 300 com 36 itens.

As populações simuladas e a curva de informação das seis provas são apresentadas na figura 1.

FIGURA 1 - Populações simuladas e curva de informação das seis provas



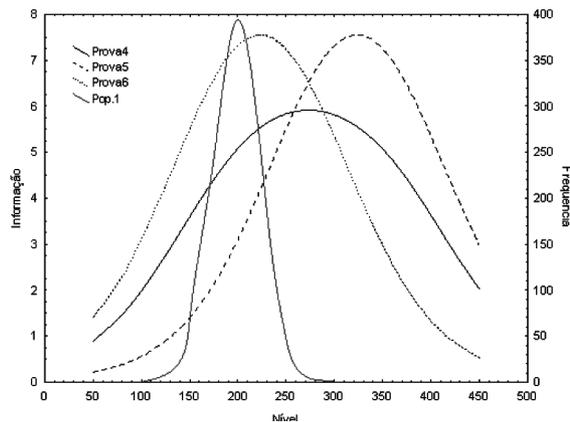
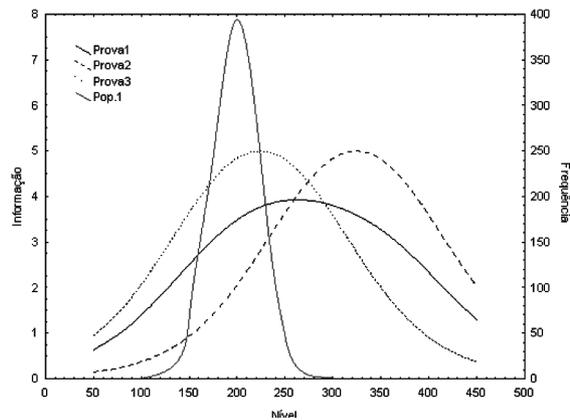
Nota-se pela figura 1 que as populações 1 e 2 são bastante distintas e homogêneas, como comentado anteriormente, já a população 3 é uma simulação que envolveria indivíduos com características das populações 1 e 2.

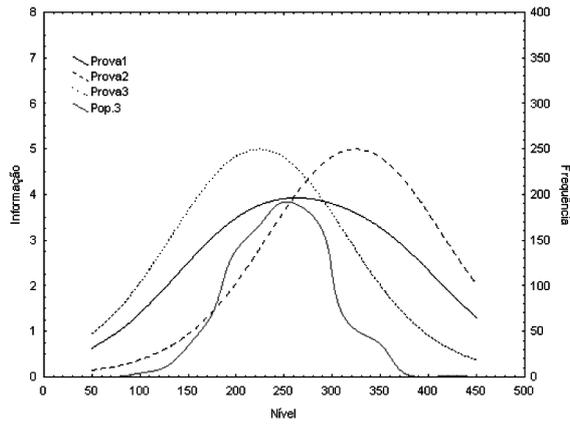
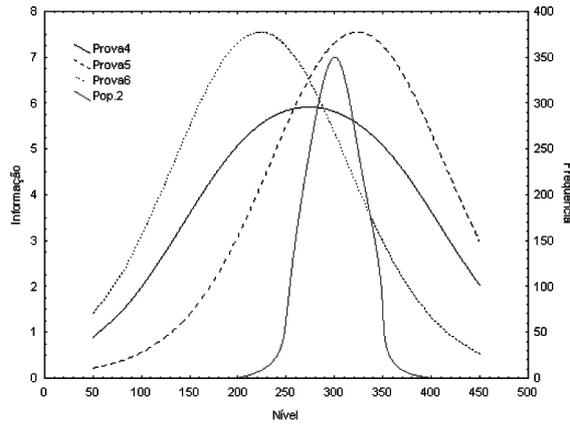
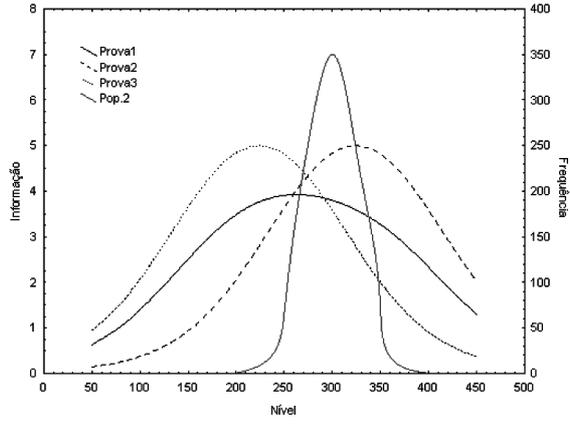
Em relação à curva de informação, nota-se que as provas 3 e 6 foram construídas para atender as características dos indivíduos da população 1, já as provas 2 e 5 para atender as características dos indivíduos da população 2, e por fim as provas 1 e 4 têm características dos indivíduos da população 3. Outra informação relevante é em relação à quantidade de itens. Nota-se que as provas com 36 itens trazem mais informação do que as provas com 24 itens, o que significa dizer que as estimativas das proficiências obtidas através das provas com 36 itens

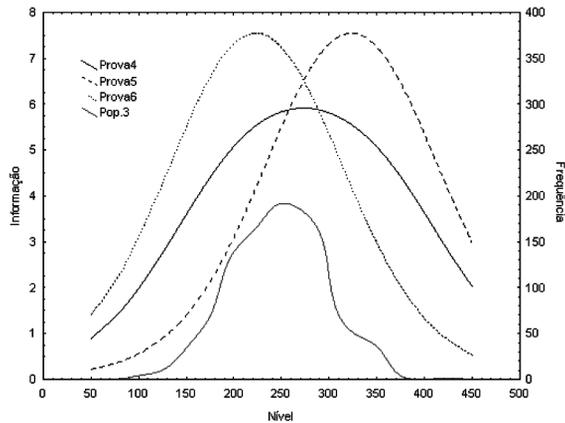
são mais precisas, na região com bastante informação, do que as provas com 24 itens.

Na figura 2 pode-se observar a região de informação das provas com 24 itens na coluna da esquerda e das provas com 36 itens na coluna da direita, e as três populações que serão submetidas a essas provas. A escala do lado direito do gráfico representa a frequência de respondentes simulados de cada população e a escala do lado esquerdo do gráfico representa a informação das curvas das provas.

FIGURA 2 - Populações e curva de informação para as provas com 24 e 36 itens







Nota-se, na figura 2, que as três populações têm as provas adequadas para suas proficiências. Como já observamos que as provas com 36 itens trazem mais informação do que as provas com 24 itens, espera-se que essas provas também forneçam melhores estimativas. Com isso, acredita-se que para a população 1 a prova 4 seja a mais adequada, já para a população 2 a prova 5 seja a melhor, e para a população 3, apesar de terem sido criadas as provas 1 e 4 com as características desta população, as provas 3 e 6 fornecem uma informação maior para respondentes com proficiência abaixo de 300.

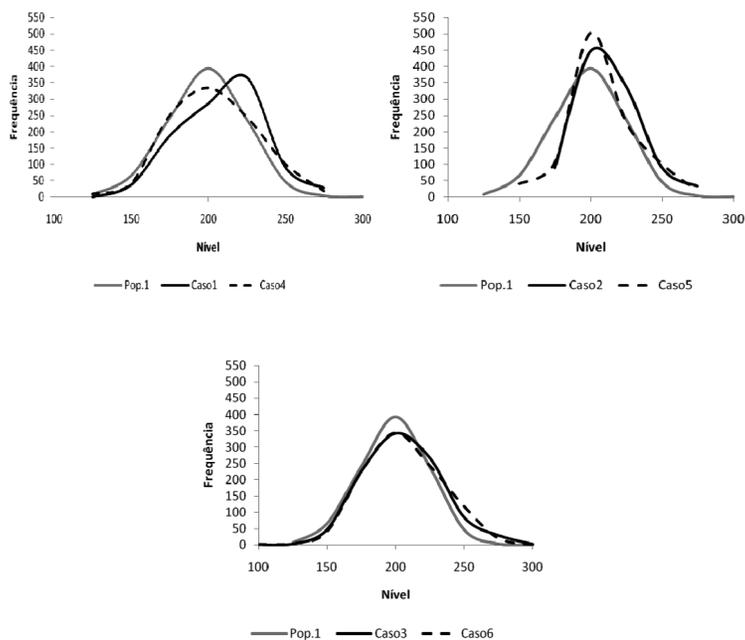
Para obter as estimativas das proficiências, cada uma das três populações foi submetida às seis provas. Com isso, temos os seguintes casos:

- | | |
|-------------------------------------|--------------------------------------|
| <i>Caso 1:</i> População 1, Prova 1 | <i>Caso 10:</i> População 2, Prova 4 |
| <i>Caso 2:</i> População 1, Prova 2 | <i>Caso 11:</i> População 2, Prova 5 |
| <i>Caso 3:</i> População 1, Prova 3 | <i>Caso 12:</i> População 2, Prova 6 |
| <i>Caso 4:</i> População 1, Prova 4 | <i>Caso 13:</i> População 3, Prova 1 |
| <i>Caso 5:</i> População 1, Prova 5 | <i>Caso 14:</i> População 3, Prova 2 |
| <i>Caso 6:</i> População 1, Prova 6 | <i>Caso 15:</i> População 3, Prova 3 |
| <i>Caso 7:</i> População 2, Prova 1 | <i>Caso 16:</i> População 3, Prova 4 |
| <i>Caso 8:</i> População 2, Prova 2 | <i>Caso 17:</i> População 3, Prova 5 |
| <i>Caso 9:</i> População 2, Prova 3 | <i>Caso 18:</i> População 3, Prova 6 |

3 ANÁLISE

Inicialmente, iremos considerar uma análise gráfica (Figura 3) das proficiências estimadas por meio das seis provas para a população 1.

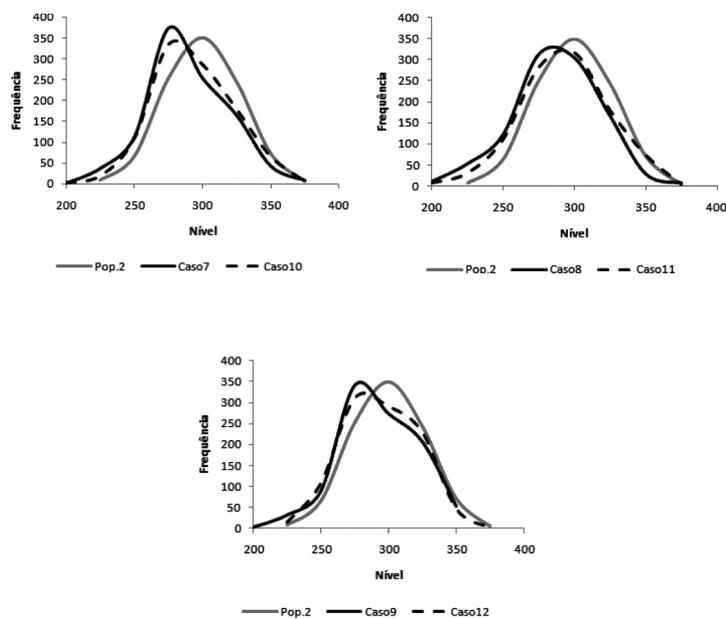
FIGURA 3 - Distribuição das proficiências verdadeira e estimada da população 1



Através da figura 3, notou-se que os casos 3 e 6, referentes às estimativas das proficiências nas provas 3 e 6, estão mais próximos da distribuição das proficiências verdadeiras. Estas provas são exatamente as provas que possuem a curva de informação próxima da distribuição das proficiências verdadeiras da população 1, conforme analisado na figura 2.

Para a população 2, as proficiências estimadas através das seis provas é dada na figura 4.

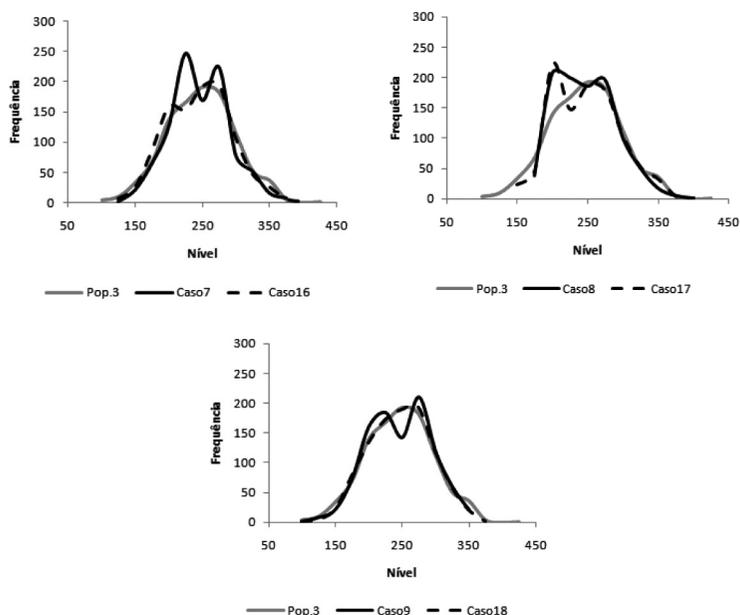
FIGURA 4 - Distribuição das proficiências verdadeira e estimada da população 2



Na figura 4, notou-se que os casos 8 e 11, referentes às proficiências estimadas para as provas 2 e 5, estão mais próximas da distribuição das proficiências verdadeiras. Estas provas são exatamente as provas que possuem a curva de informação próxima da distribuição das proficiências verdadeiras da população 2, conforme analisado na figura 2. Entretanto, observa-se que em todas as provas, as estimativas das proficiências subestimam um pouco as proficiências verdadeiras.

Para a população 3, as proficiências estimadas através das seis provas é dada na figura 5.

FIGURA 5 - Distribuição das proficiências verdadeira e estimada da população 3



Nota-se, na figura 5, que as estimativas das proficiências são mais espalhadas do que para as duas primeiras populações. Isto ocorreu em razão da maior variabilidade na simulação desta população. O caso 18, referente às proficiências estimadas para a prova 6, apresenta a distribuição estimada mais próxima da distribuição verdadeira.

A tabela 1 fornece informações do Erro Quadrático Médio (EQM) para as seis provas, considerando as três populações.

TABELA 1 - Erro Quadrático Médio para as três populações

	POPULAÇÃO 1	POPULAÇÃO 2	POPULAÇÃO 3
Prova 1	560,33	561,45	540,76
Prova 2	686,93	564,80	623,90
Prova 3	430,47	579,99	460,73
Prova 4	399,04	401,18	377,42
Prova 5	518,51	390,88	469,87
Prova 6	306,46	397,03	336,22

Conforme os resultados analisados na distribuição das proficiências, a prova 6 está gerando um erro menor para as populações 1 e 3 e para a população 2 o erro foi menor para a prova 5. Além disso, os maiores erros em cada população ocorreram para as provas com 24 itens, evidenciando que o aumento de item resultará em melhores estimativas das proficiências.

Para confirmar os resultados acima, a tabela 2 apresenta a correlação das proficiências estimadas com as verdadeiras proficiências de cada população, observando que as maiores correlações ocorrem para a população 3.

TABELA 2 – Correlações entre as estimativas das proficiências e as proficiências verdadeiras

	POPULAÇÃO 1	POPULAÇÃO 2	POPULAÇÃO 3
Prova 1	560,33	561,45	540,76
Prova 2	686,93	564,80	623,90
Prova 3	430,47	579,99	460,73
Prova 4	399,04	401,18	377,42
Prova 5	518,51	390,88	469,87
Prova 6	306,46	397,03	336,22

4 CONCLUSÕES

Este trabalho representa um primeiro estudo sobre a influência da prova (teste) na estimativa da proficiência de um aluno, através da Teoria da Resposta ao Item. Foi mostrado que um teste com informação na região da proficiência do respondente fornece estimativas mais precisas do que testes com informações em outras regiões da proficiência. Além disso, mostrou-se também a importância de testes maiores na estimação da proficiência.

Novos estudos serão realizados considerando também os modelos de dois e três parâmetros da TRI.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, D. F., TAVARES, H. R., VALLE, R. C. *Teoria da Resposta ao Item: conceitos e aplicações*. Caxambu: Associação Brasileira de Estatística, 2000.

AZEVEDO, C. L. N. *Métodos de estimação na Teoria da Resposta ao Item*.
Dissertação de Mestrado. IME/USP, 2003.

BAKER, F. B.; KIM, S. H. *Item Response Theory: Parameter Estimation
Techniques*. New York: Marcel Dekker, Inc., 2004.

CONDE, F. N. A *(In)Dependência da habilidade estimada pela Teoria de
Resposta ao Item em relação à dificuldade da prova: um estudo com
dados do Saeb*. Dissertação de Mestrado em Psicologia na UnB, 2002.

LORD, F. M. *Applications of Item Response Theory to Practical Testing
Problems*. Hillsdale: Lawrence Erlbaum Associates, Inc., 1980.

DALTON FRANCISCO DE ANDRADE

Professor voluntário junto ao Programa de Pós-Graduação
do Departamento de Engenharia de Produção da
Universidade Federal de Santa Catarina
dandrade@inf.ufsc.br

ADRIANO FERRETI BORGATTO

Professor adjunto do departamento de Informática
e Estatística/Centro Tecnológico (INE/CTC) da
Universidade Federal de Santa Catarina
borgatto@inf.ufsc.br