

## Equalização de provas acadêmicas via Teoria de Resposta ao Item

Fernanda Luzia Lopes – Universidade São Francisco, Itatiba, Brasil  
 Claudette Maria Medeiros Vendramini – Universidade São Francisco, Itatiba, Brasil

---

### Resumo

Este estudo objetivou equalizar uma prova acadêmica de Pedagogia com o ENADE via Teoria de Resposta ao Item. Utilizou-se um banco de dados contendo informações de 49.497 estudantes de Pedagogia que realizaram o ENADE em 2005 e um banco de dados contendo informações de 260 estudantes de uma Instituição de Ensino Superior – IES que realizaram uma prova de Pedagogia de um Programa de Avaliação Discente – PAD. Os resultados apontaram que o desempenho dos estudantes concluintes da IES na prova PAD obteve notas médias percentilicas semelhantes ao desempenho dos estudantes com conceito 2 que realizaram o ENADE em 2005. Os concluintes da IES que realizaram o PAD alcançaram notas médias semelhantes às notas médias dos concluintes com conceito 4 que realizaram o ENADE. Este trabalho é um pequeno recorte alusivo às teorias e técnicas de análise de dados, com o intuito de colaborar com esse contexto de estudo ainda em construção.

*Palavras-chave:* Avaliação em larga escala, Métodos estatísticos, Ensino superior, Psicometria.

### Equalization of academic tests via Item Response Theory

#### Abstract

This study aimed to equalize a Pedagogical academic test based on National Performance Student Exam via Item Response Theory. We used a database containing information of 49,497 Pedagogy students that took the exam in 2005 and another database containing information of 206 students of a higher education institution, who took a Pedagogy test of a Student Assessment Program – PAD. The results indicated that de senior students' performance of the higher educational institution in the PAD's test presented average percentilic grades similar to the student's performance with concept 2 who took the exam in 2005. The senior students of the higher educational institution who took the PAD's test reached average grades similar to the grade of the senior students who obtained concept 2 in 2005. This work is a small sample alluding to equalization theories and techniques in order to collaborate to this study context.

Keywords: Large scale assessment, Statistical methods, Higher education, Psychometrics.

### Ecualización de pruebas académicas vía Teoría de Respuesta al Ítem

#### Resumen

Este estudio objetivó ecualizar una prueba académica de Pedagogía con el ENADE vía Teoría de la Respuesta al Ítem. Se utilizó una base de datos que contiene información de 49.497 estudiantes de Pedagogía que realizaron el ENADE en 2005 y una base de datos de 260 estudiantes de una Institución de Educación Superior – IES que realizaron una prueba de un Programa de Evaluación Discente – PAD. Los resultados mostraron que el rendimiento de los estudiantes de final de carrera de la IES en la prueba PAD alcanzó calificaciones medias percentilicas semejantes al rendimiento de los Estudiantes con concepto 2 que realizaron el ENADE. Los estudiantes del final de carrera de la IES que realizaron la prueba PAD alcanzaron calificaciones medias semejantes a los estudiantes de final de carrera con concepto 4 en el ENADE. Este trabajo es un pequeño recorte alusivo sobre las teorías y las técnicas de análisis de datos, con el objetivo de colaborar con ese contexto de estudio aún en construcción.

*Palabras-clave:* Evaluación a larga escala, Métodos estadísticos, Educación superior, Psicometría.

A preocupação com a qualidade do ensino ofertado não é somente incumbência do governo federal. Cada vez mais as instituições de ensino superior, motivadas pelo Sistema Nacional de Avaliação do Ensino Superior (SINAES) instituído no ano de 2004 e tendo como uma de suas ferramentas o Exame Nacional de Desempenho dos Estudantes (ENADE), vêm zelando pelos atributos de suas graduações e, assim, criando seus próprios sistemas de avaliação. Não diferente, essas avaliações também necessitam dos rigores científicos para que sejam consideradas a integridade e a qualidade dos instrumentos adotados. Assim, os métodos psicométricos podem auxiliar na análise e verificação de evidências de validade e fidedignidade dos

instrumentos utilizados para a avaliação do ensino e da aprendizagem.

Tendo em vista a preocupação com a qualidade dos cursos de graduação e o progresso de seus estudantes, no ano de 2008 foi instituído o Programa de Avaliação Discente (PAD) em uma universidade particular do interior paulista, com edições anuais, que contemplam, a cada ano, distintas áreas e cursos de graduação. O PAD é um sistema de avaliação institucional que conta com provas construídas de acordo com a matriz do ENADE.

O PAD é composto por 40 itens, sendo 10 de formação geral e 30 de conhecimento específico, nos mesmos moldes da prova do ENADE. Alguns itens são selecionados entre os de melhor qualidade psicométrica do ENADE e servirão como itens âncora

na equalização das provas ENADE e PAD. A partir da equalização dessas provas é possível comparar o desempenho dos estudantes da Instituição de Ensino Superior (IES) em relação aos estudantes brasileiros dos respectivos cursos analisados.

Esse modelo de avaliação interna se torna possível com a utilização de recursos técnicos obtidos pela equalização de provas, uma das ferramentas da Teoria de Resposta ao Item (TRI), que permite a criação de uma escala de referência em que as notas dos estudantes possam ser passíveis de comparação. Com isso, além da comparação entre os desempenhos acadêmicos dos estudantes, também é plausível a investigação do progresso desses estudantes durante os diferentes períodos da graduação, não somente demonstrando sua importância para a questão da qualidade dos cursos oferecidos, mas também servindo de instrumento para diagnósticos acadêmicos que possam servir de base para futuras tomadas de decisões.

Nos últimos 20 anos, a TRI vem sendo amplamente difundida como um sistema de medição que se faz pertinente em diversas áreas e saberes e, principalmente, na avaliação educacional (Andrade, Tavares & Valle, 2000; Pasquali, 2007). Segundo Conde e Laros (2007), o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) vem utilizando a TRI desde 1995; já que, por se tratar de uma avaliação realizada em caráter nacional, ou seja, em larga escala, a metodologia dessa teoria possibilita o melhor conhecimento do desempenho dos estudantes e das variáveis suscetíveis a esse contexto.

Por se tratar de um modelo estatístico, a TRI permite calcular a probabilidade de um sujeito acertar um determinado item e, ainda, estimar a habilidade latente necessária para a resolução de um problema (Andrade & Valle, 1998; Conde & Laros, 2007; Pasquali & Primi, 2007; Tejada & Meléndez, 2001; Valle, 2000; Vendramini, 2005;). No entanto, a habilidade latente não pode ser mensurada diretamente; Conde e Laros (2007, p. 205) explicam que “na TRI a probabilidade de acertar o item é modelada como função da habilidade latente do indivíduo e os parâmetros que representam algumas propriedades do item”.

As habilidades (traços latentes) consideradas necessárias para o sucesso na resposta a um determinado item não são passíveis de observação direta, mas podem ser objeto de inferências por meio do comportamento do sujeito diante de problemas ou situações apresentados com essa finalidade. Dessa maneira, a TRI considera que as estimativas dos traços latentes são dependentes das respostas dos sujeitos e

das propriedades componentes do item (Embretson & Reise, 2000).

Assim, a relação aparente entre variáveis hipotéticas (traços latentes) e variáveis observadas pode ser expressa por meio de um modelo matemático. Por exemplo, os itens de um teste podem ser considerados variáveis observadas e, quando se conhecem suas características, estas se tornam constantes em uma equação e, dessa maneira, solucionáveis. A partir desse ponto, torna-se possível estimar o nível do traço latente e, assim, compreender as características dos itens respondidos pelo sujeito (Andrade & Valle, 1998; Valle, 2000; Pasquali & Primi, 2007; Vendramini, 2005).

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

A equalização é uma importante ferramenta quando o objetivo do pesquisador educacional é comparar o desempenho de indivíduos pertencentes ou não à mesma população. Para Andrade, Tavares e Valle (2000), a equalização permite a comparação entre itens e/ou níveis de habilidades, já que essa ferramenta possibilita que os parâmetros de itens e/ou níveis de habilidades dos respondentes, de grupos diferentes, possam ser analisados e compreendidos em uma mesma métrica, ou seja, uma escala comum.

Conforme Araújo, Andrade e Bortolotti (2009, p. 1006),

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

Ampliando a definição, Silva e Soares (2010) destacam que a equalização refere-se à comparação de resultados, de formas distintas, de um mesmo teste projetado para ser paralelo. Dessa maneira,

*os testes medem os mesmos conteúdos, possuem os mesmos descritores, a mesma estrutura, mesma forma de aplicação, pequena variação na dificuldade de itens similares que compõem as diferentes formas dos testes e*

*as populações equivalentes.* (Silva & Soares, 2010, p. 194)

Tais características permitem resultados eficazes, garantindo, assim, para as diferentes formas, o mesmo nível de confiabilidade.

Diferentes tipos de equalização podem ser usados de acordo com os propósitos e ferramentas disponíveis ao pesquisador, a saber, um único grupo fazendo duas provas totalmente distintas, conhecida como equalização via população – “o fato de todos os indivíduos representarem uma amostra aleatória de uma mesma população é que garante que todos os parâmetros envolvidos estarão numa mesma escala” (Santos, 2008, p. 43) –; um único grupo fazendo duas provas parcialmente distintas, semelhante ao caso anterior; dois grupos fazendo uma única prova, com a equalização via itens comuns em que se calibram os itens utilizando-se as respostas dos participantes de ambos os grupos, simultaneamente; e, por fim, dois grupos fazendo duas provas parcialmente distintas (Santos, 2008).

Esse último tipo de equalização, conhecida também como equalização via itens comuns, utiliza-se de itens âncoras, ou seja, itens que são parte de uma prova e servirão, de acordo com suas características psicométricas, como itens de referência entre as provas.

*O uso de itens comuns entre provas distintas aplicadas a populações distintas permite que todos os parâmetros estejam na mesma escala ao final dos processos de estimação, possibilitando comparações e a construção de “escalas de conhecimento” interpretáveis que são de grande importância na área educacional.* (Santos, 2008, p. 43)

Contribuindo com o exposto, Silva e cols. (2011) destacam que a equalização via TRI possibilita que diferentes formas de um mesmo teste ou testes distintos, mas que contenham itens comuns, possam ser passíveis de comparação. Assim, a partir da estimação dos parâmetros de cada item que servirá de âncora à prova e com as informações acerca do nível de dificuldade e, ainda, que mencionem a capacidade do estudante necessária à execução eficiente do item, é possível determinar ligação entre as populações abrangidas e permitir que os parâmetros sejam estimados em uma escala comum, ou seja, que seus desempenhos sejam referenciados da mesma forma. Cabe ressaltar que o procedimento descrito pelos autores é o mesmo que embasa a construção e as análises das provas PAD.

Pesquisas que fazem uso da equalização na situação de Avaliação da Educação Superior ainda são escassas. Nesse sentido, destacam-se os estudos desenvolvidos pelo LabAPE (Laboratório de Avaliação Psicológica e Educacional), vinculado à Universidade

São Francisco, com os dados do ENADE, como os de Primi e cols. (2010), Silva e cols. (2011) e Bartholomeu e cols. (2011). Considerando a equalização como um instrumento que possibilita estudos que visem a qualidade da avaliação da educação superior, para o presente estudo o objetivo foi de equalizar provas de habilidades acadêmicas, a partir do ENADE, via TRI.

## Método

### *Fonte de dados*

Os participantes são parte de um banco de dados que contém os resultados dos estudantes do curso de Pedagogia, que realizaram a prova do ENADE no ano de 2005. Ressalta-se que esta pesquisa está vinculada a um projeto maior, a saber, “A Validade do ENADE para a Avaliação da Qualidade dos Cursos de Instituições do Ensino Superior”, conforme edital estabelecido pelo MEC para a instituição de Observatórios da Educação em IES, apoiados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior do Ministério da Educação (Capes/MEC). O Observatório da Educação, particularmente o da Universidade São Francisco, tem por objetivos prioritários buscar evidências de validade da estrutura interna dos itens do ENADE, compreender o sentido das dimensões encontradas empregando métodos da Teoria de Resposta ao Item (TRI) e, ainda, investigar as evidências de validade por intermédio da verificação das relações com outras variáveis (Primi, 2006).

Para tanto, o projeto conta com os bancos de dados nacionais do ENADE, concedidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), órgão atrelado ao Ministério da Educação. Os bancos de dados são compostos por estudantes selecionados aleatoriamente por amostragem estratificada, de acordo com os critérios do MEC, e contam com informações referentes ao curso, às IES, respostas ao componente de formação geral, ao componente específico por curso e ao questionário socioeconômico. Este estudo também contou com um banco de dados contendo informações referentes às respostas dos estudantes do curso de Pedagogia a uma Prova PAD equalizada ao ENADE 2005. Essas respostas contemplam dois componentes, quais sejam a Prova de Formação Geral e a prova de conhecimento específico, totalizando 40 questões.

### *Participantes*

Participaram do ENADE de Pedagogia, no ano de 2005, 49.497 estudantes de todos os estados brasileiros e Distrito Federal, referindo-se a maior concentração à região sudeste do país, 49,4%. Os estudantes, com idade média de 30,05 anos ( $DP=8,88$ ),

foram divididos em 52,9% ingressantes e 47,1% concluintes. A maioria dos ingressantes (92,6%) é do sexo feminino, assim como dos concluintes (92,0%).

Também foram participantes 260 estudantes matriculados no curso de pedagogia de uma universidade particular do interior do estado de São Paulo, em 2010, dos quais 93,1% mulheres, em diferentes momentos do curso, distribuídos em três grupos, a saber, ingressantes, intermediários e concluintes (Tabela 1). Dessa forma, estudantes que cursaram até 25% (1º e 2º semestres) do curso serão considerados ingressantes; aqueles que cumpriram até

75% (3º ao 6º semestres) da graduação serão observados como intermediários e os últimos, acima de 75% (7º e 8º semestres) do curso realizado, considerados como concluintes.

Esse critério para a distribuição e agrupamento dos estudantes é semelhante aos critérios utilizados pelo MEC, diferindo, apenas, no que concerne aos estudantes considerados intermediários, já que essa faixa dos estudantes não é avaliada pelo ENADE. A Tabela 1 mostra a distribuição de estudantes por diferenciação acadêmica, em que a maioria de 59,6% é considerada intermediária.

Tabela 1. Distribuição dos estudantes pela situação acadêmica

Situação acadêmica	Quantidade de estudantes	
	Nº	%
Ingressantes	42	16,2
Intermediários	155	59,6
Concluintes	63	24,2
Total	260	100,0

### Material

Serviu de material uma prova do PAD de pedagogia, constituída com o formato semelhante ao ENADE, composta por 40 itens subdivididos em, 10 itens de formação geral, comuns a todos os cursos, 5 deles comuns ao ENADE de pedagogia de 2005; 10 itens de conhecimento específico referentes ao Núcleo de Conteúdos Básicos ou Fundamentais; 20 itens de conhecimento específico referentes ao Núcleo de Conteúdos Profissionalizantes dos quais 11 são comuns ao ENADE de pedagogia de 2005.

### Procedimento de análise dos dados

Foi solicitada a autorização junto à Instituição de Ensino Superior para o uso da base de dados proveniente da aplicação dessas provas. Os dados foram organizados a fim de se atender às necessidades dos *softwares* utilizados na pesquisa, em que as frequências de acertos dos estudantes, tanto do ENADE 2005 quanto do PAD, foram transformadas em percentis, a fim de uma melhor análise dos resultados, analisar as propriedades psicométricas dos itens que compõem a prova de pedagogia do ENADE 2005. Para tanto, foram seguidas as seguintes etapas: seleção de itens que apresentaram boas propriedades psicométricas para compor a prova PAD de pedagogia; composição da prova PAD com os itens selecionados do ENADE e outros elaborados pelos docentes do curso de pedagogia da IES; aplicação da prova PAD aos estudantes da IES; cálculo do nível de habilidade dos estudantes e do nível de dificuldade dos itens, por meio da equiparação via TRI entre o ENADE e a prova PAD, representados em uma mesma escala de

medida; transformação do nível de habilidade dos estudantes, calculados via TRI, em percentis que indicam a porcentagem de estudantes com valores menores ou iguais à ordem percentilica de cada medida de habilidade; verificação das possíveis diferenças no desempenho dos estudantes que fizeram a prova PAD de pedagogia e os que fizeram o ENADE 2005 de pedagogia, por meio da representação gráfica do nível de habilidade transformado em nota média percentilica.

### Resultados e Discussão

O presente estudo teve por objetivo apresentar os resultados da equalização via TRI da prova PAD de conhecimento específico, composta de 30 itens, aplicada ao curso de pedagogia em 2010, comparada ao ENADE 2005. Os resultados dessa equalização permitiram verificar a classificação do nível de habilidade de cada estudante avaliado, tanto em relação aos estudantes brasileiros quanto em relação aos estudantes da IES.

Com a finalidade de verificar se o nível de habilidade dos estudantes aumenta com a melhor classificação dos cursos (conceitos) realizada pelo INEP, inicialmente os estudantes foram agrupados em ingressantes e concluintes, de acordo com esses conceitos atribuídos pelo INEP (2005) aos cursos de pedagogia vigentes no país. Os conceitos de cada curso podem variar de 1 a 5 e correspondem à nota média padronizada dos estudantes nos componente de formação geral (25% da nota total) e conhecimento específico (75% da nota total). A Tabela 2 aporta a distribuição dos estudantes ingressantes e concluintes

de pedagogia do Brasil por conceito ENADE 2005. Por meio desta Tabela é possível visualizar que a maior parte dos estudantes, tanto os ingressantes (49,5%) quanto os concluintes (51,1%), foram categorizados no conceito 3 do ENADE, totalizando 50,3% de toda a

amostra. Cabe ressaltar que os estudantes considerados ingressantes cursavam, na ocasião do exame, do 1º ao 3º semestre do curso, e os considerados ingressantes, os 7º e 8º semestres da graduação (INEP, 2005).

Tabela 2. Distribuição de estudantes de pedagogia do Brasil por conceito no ENADE 2005

Conceito	Ingressantes		Concluintes		Total	
	Nº	%	Nº	%	Nº	%
1	609	2,6	706	2,7	1.315	2,7
2	2.785	11,9	3.515	13,4	6.300	12,7
3	11.552	49,5	13.366	51,1	24.918	50,3
4	4.737	20,3	5.018	19,2	9.755	19,7
5	301	1,3	293	1,1	594	1,2
SC	3.334	14,3	3.281	12,5	6.615	13,4
Total	23.318	100,0	26.179	100,0	49.497	100,0

Nota: SC = Sem Conceito

Da mesma forma, espera-se que estudantes matriculados em séries mais avançadas apresentem nível de habilidade mais elevado. Portanto, os estudantes que fizeram o PAD foram agrupados por ano de curso em que estão matriculados, correspondendo a categoria ingressante do ENADE aos estudantes do 1º ano do PAD, equivalente a 16,2% do total de alunos; as categorias 2º (40,5%) e 3º (19,3%) anos são de estudantes intermediários, sem correspondência ao ENADE; e, por fim, os estudantes do 4º ano (23,9%) equivalem à categoria Concluintes do ENADE. Os estudantes do PAD também foram agrupados por campus em que estavam matriculados, concentrando-se a maior parte deles no Campus 2 da IES (57,9%).

A partir da equalização das provas PAD e ENADE 2005 de pedagogia, foi possível comparar o desempenho dos estudantes por meio de uma escala de medida comum, em que o sem número de acertos é equivalente a determinados níveis de habilidade. Esses

níveis são transformados em uma escala de percentis que pode variar de 1 a 99 e indica a proporção de estudantes do grupo de referência que obtiveram a mesma quantidade ou menos acertos que a nota do estudante em foco (PAD). Assim, um estudante cuja nota equivale ao Percentil 50 tem uma nota que supera, aproximadamente, 50% das notas do grupo de referência.

Cabe salientar que, em virtude da equalização das notas PAD com os do ENADE, é possível ter como grupo de referência todos os estudantes do Brasil que fizeram o ENADE em 2005 – as notas percentílicas têm como grupo de referência a amostra brasileira do exame. Assim sendo, os percentis foram calculados a partir das notas do ENADE e, com essa equalização, os valores referentes à escala PAD correspondem ao mesmo nível de habilidade no ENADE. A distribuição dos percentis da nota média no ENADE, agrupados em ingressantes e concluintes, é apresentada na Figura 1.

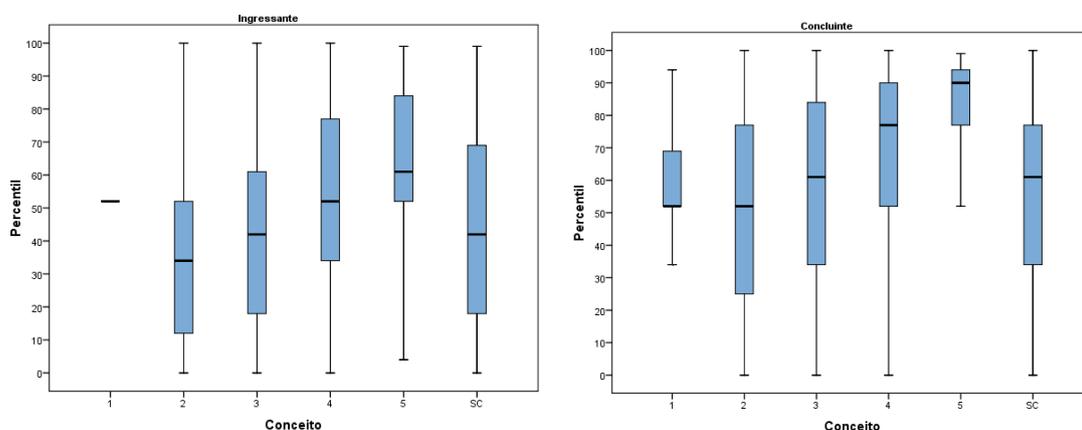


Figura 1. Diagrama de caixas das notas médias percentílicas no ENADE de ingressantes e concluintes por conceito do curso

A análise da Figura 1 permite dizer que, de acordo com os percentis obtidos, o desempenho dos estudantes que participaram do ENADE de pedagogia em 2005 aumenta conforme aumenta o conceito dos cursos, resultados esperados, ou seja, espera-se um melhor desempenho dos estudantes cujos cursos obtiveram conceitos mais altos. Dessa forma, tanto os estudantes ingressantes quanto os concluintes obtiveram melhor desempenho nos Cursos com conceito 5 no ENADE. Salienta-se que, os cursos que não contaram com a presença de ingressantes ou concluintes foram classificados como Sem Conceito (SC), haja vista a indisponibilidade do cálculo das notas finais desses estudantes.

A Figura 2 contempla os percentis de nota média dos estudantes do PAD, agrupados por ano e campus

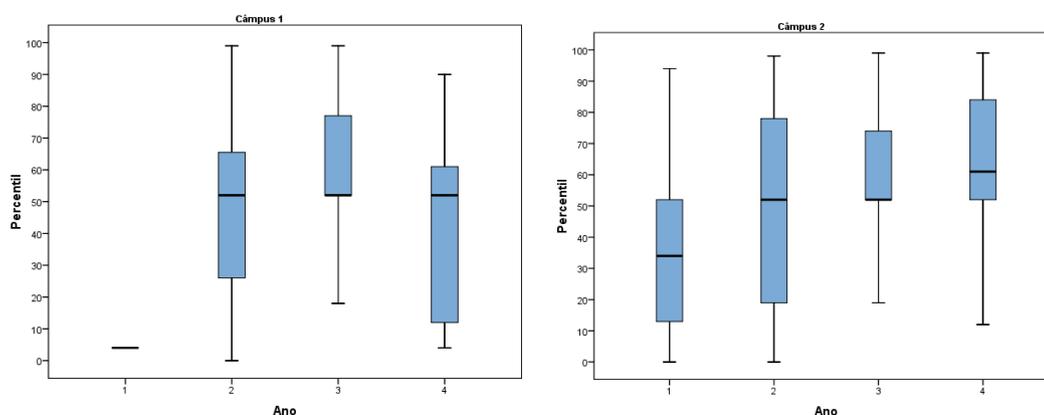


Figura 2. Diagrama de caixas das notas médias percentílicas no PAD por ano e campus.

Já para os estudantes do campus 2, houve uma variabilidade consonante aos resultados esperados, ou seja, espera-se um melhor desempenho dos estudantes conforme vão avançando na graduação. Tomando como exemplo os estudantes do 3º ano, pôde-se observar que a nota mínima destes corresponde ao percentil 20, a mediana está em torno do percentil 52 e a nota máxima corresponde, aproximadamente, ao percentil 99. Os estudantes do 4º ano angariaram a nota mínima que satisfaz o percentil de, aproximadamente, 15; a mediana foi por volta do percentil 62 e, por último, a nota máxima dos estudantes ficou em torno do percentil 99. As notas dos dois anos de graduação foram semelhantes, mas, como pode ser visto na Figura 2, pelo menos 50% dos alunos matriculados no 4º ano obtiveram melhor desempenho no PAD.

Na comparação do desempenho entre os estudantes que fizeram o ENADE e o PAD, não puderam ser utilizadas as notas dos alunos ingressantes

de matrícula. Conforme a figura pode-se observar variabilidade no desempenho dos estudantes de acordo com a seqüência dos anos de graduação; contudo, deve-se enfatizar que os estudantes do campus 1, matriculados no 3º ano de graduação mostraram melhor desempenho que os estudantes matriculados no 4º ano do curso. As medidas aproximadas referentes ao desempenho dos estudantes do 3º ano do PAD indicaram nota mínima correspondente ao percentil 20, mediana em torno do percentil 52, e a nota máxima alude ao percentil 99. Essas mesmas medidas para os estudantes do 4º ano denotam nota mínima congruente com o percentil 5, mediana em torno do percentil 52, e a nota máxima por volta do percentil 90.

do PAD (matriculados no 1º ano) dos cursos do campus 1 devido ao pequeno número de participantes. Já com relação ao desempenho dos alunos concluintes do PAD, matriculados no 4º semestre da graduação do campus 1, em comparação aos alunos concluintes do ENADE, os primeiros obtiveram notas médias percentílicas semelhantes ao desempenho dos alunos com conceito 2 no ENADE.

A análise do desempenho dos estudantes ingressantes do PAD matriculados no campus 2, quando equalizados à escala do ENADE, identificou que, de acordo com as figuras 1 e 2, esses alunos tiveram desempenho semelhante aos estudantes ingressantes dos cursos com Conceito 2 no ENADE. Considerando o desempenho dos estudantes concluintes do PAD, ou seja, matriculados no último ano de pedagogia do campus 2, segundo as figuras 1 e 2, esses alcançaram notas médias semelhantes às notas médias alcançadas pelos estudantes concluintes dos cursos com conceito 4 no ENADE, haja vista o extrato

de 50% de estudantes com notas entre percentil 25 e 75.

Levando-se em conta os dois campi que fizeram o PAD de pedagogia, observa-se que os estudantes considerados intermediários que fizeram a prova nos dois campi apresentam desempenhos semelhantes, para os estudantes de 4º ano do campus 2, a nota aumenta em relação aos anos anteriores e a nota dos estudantes do campus 1 permanece estável. Pelo pequeno número de estudantes ingressantes que realizaram a prova no campus 1 não é possível fazer comparações entre estudantes ingressantes dos dois cursos.

Os processos de equalização de provas de habilidades acadêmicas que buscam verificar/equiparar o desempenho de estudantes ainda são escassos. Na literatura especializada, estão descritos esses processos, principalmente, com dados provenientes do Sistema Nacional de Avaliação da Educação Básica – SAEB. Nesse sentido, Locatelli (2002) informa que os dados do SAEB, desde a década de 1990, vêm se utilizando da TRI como metodologia de equalização e análise de dados. Assim, calculados os parâmetros de dificuldade, discriminação e o acerto ao acaso, os itens considerados com boas propriedades psicométricas são armazenados em um banco nacional de itens e podem ser usados na situação de avaliações nacionais e estaduais. A autora enfatiza as vantagens da equalização quando se pretende colocar os resultados dos testes em uma mesma escala, possibilitando a comparação entre diferentes grupos de estudantes. Para obter maiores informações sobre a equalização com os dados do SAEB, ver estudos de Klein (2009) e Souza (2005).

Na avaliação do Ensino Superior que emprega o uso da equalização via TRI, destacam-se os estudos realizados pelo LabAPE (Laboratório de Avaliação Psicológica e Educacional), vinculado ao Programa de Pós-Graduação *Stricto Sensu* da Universidade São Francisco. Dentre as pesquisas realizadas pelo LabAPE está a realização de provas institucionais construídas de modo semelhante à metodologia utilizada nesta pesquisa desde a instituição do PAD, no ano de 2008. Os resultados desses procedimentos são descritos em relatórios técnicos que contemplam as análises psicométricas das provas do ENADE, a seleção dos itens que compõem as provas PAD, as análises psicométricas das provas PAD e, por fim, a comparação de desempenhos entre estudantes PAD e ENADE (Primi & cols., 2010). Além dessas produções, pesquisas como as de Silva e cols. (2011) e de Bartholomeu e cols. (2011) demonstram a abrangência das análises e algumas aplicações que podem ser feitas por meio do uso da equalização e TRI.

## Considerações finais

O presente estudo teve por objetivo equalizar uma prova de habilidades acadêmicas utilizando-se do Exame Nacional de Desempenho dos Estudantes com o auxílio dos modelos matemáticos propostos pela Teoria de Resposta ao Item. Pôde-se averiguar que o modelo matemático proposto pela TRI mostrou-se bem sucedido para a equalização de provas de habilidades acadêmicas. O produto da equalização propriamente dita permitiu a comparação entre os estudantes que fizeram a prova PAD de Pedagogia e aqueles que fizeram a prova de pedagogia do ENADE no ano de 2005. Esses resultados se fazem pertinentes em relação aos objetivos estipulados para o PAD e mostram a eficácia da equalização por meio da TRI.

Diante de todo o exposto, este estudo não teve o propósito de extenuar o assunto; pelo contrário, tendo em vista que muito ainda se pode construir na direção de agregar técnicas, metodologias, percepções e construções teóricas no que se refere à interface Avaliação da Educação Superior e Teoria de Resposta ao Item. Nesse sentido, o presente trabalho mostra-se, somente, como um pequeno recorte alusivo sobre as teorias e as técnicas possíveis, com o intuito de corroborar, ao menos em partes, esse contexto ainda em construção.

## Referências

- Andrade, D. F., & Valle, R. C. (1998). Introdução à Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, 18, 13-32.
- Andrade, D. F., Tavares, H. R., & Valle, R. C. (2000). *Teoria de Resposta ao Item: conceitos e aplicações*. São Paulo: ABE.
- Araújo, E. A. C., Andrade, D. F. de, & Bortolotti, S. L. V. (2009). Teoria de Resposta ao Item. *Revista da Escola de Enfermagem da USP*, 43(n.esp), 1000-1008.
- Bartholomeu, D., Silva, M. C. R., Pires, S. D., & Primi, R. (2011). Estabelecimento de pontos de corte pela Teoria de Resposta ao Item (TRI) para uma prova equalizada com o ENADE. Em Fernando C. Capovilla (Org.). *Transtorno de aprendizagem: progressos em avaliação e intervenção preventiva e remediativa* (pp. 303-312). São Paulo: Click Books.
- Conde, F. N., & Laros, A. J. (2007). Unidimensionalidade e a propriedade de invariância das estimativas da habilidade pela TRI. *Avaliação Psicológica* 6(2), 205-215.

- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum.
- Klein, R. (2009). Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB). *Meta: Avaliação*, 1(2), 124-140.
- Locatelli, I. (2002). O Sistema de Avaliação da Educação Brasileira. Em *La experiencia brasileña em la Educacion Básica. IV Simposio Iberoamericano de Investigación y Educación, I Simposio Iberoamericano de Investigación sobre Eficacia Escolar y Mejora de la Escuela*, Habana, Cuba.
- Pasquali, L. (2007). *Teoria de Resposta ao Item – TRI*. Brasília, DF: LabPAM/UnB.
- Pasquali, L., & Primi, R. (2007). Fundamentos da Teoria da Resposta ao Item – TRI. Em L. Pasquali (Org.). *Teoria de Resposta ao Item: TRI* (pp. 11-28). Brasília, Distrito Federal: Editora UNB.
- Primi, R. (2006). *A validade do ENADE para avaliação da qualidade dos cursos de instituições de ensino superior*. Projeto de Pesquisa. Itatiba: Universidade São Francisco, LabAPE
- Primi, R., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Carvalho, L. F., Miguel, F. K., & Vendramini, C. M. M. (2009a). Aplicação da Teoria de Resposta ao Item na interpretação das notas do ENADE de Psicologia. *Revista de Educação ANEC*, 38(150), 115-124.
- Primi, R., Silva, M. C. R., Bartholomeu, D., Vendramini, C. M. M., & Silva, C. H. S. S. (2009b). Questões metodológicas referentes ao Exame Nacional de Desempenho dos Estudantes (ENADE). *Revista de Educação ANEC*, 38(150), 125-134.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Lopes F. L., Marietto, R. M., Pires, S. D., Lima A. L., Silva, M. C. R., & Bartholomeu, D. (2010). *Análise psicométrica e resultados dos alunos da USF pedagogia*. Relatório Técnico Institucional não publicado.
- Santos, L. M. (2008). *Desempenho escolar em Pernambuco: análise dos itens e das habilidades usando a Teoria Clássica e TRI*. Dissertação de Mestrado, Departamento de Estatística, Universidade Federal de Pernambuco, Recife, PE.
- Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Primi, R. (2011). Teoria de Resposta ao Item: equalização da prova de administração do ENADE. Em Fernando C. Capovilla. *Transtorno de aprendizagem: progressos em avaliação e intervenção preventiva e remediativa* (pp. 313-320). São Paulo: Click Books.
- Silva, W., & Soares, T. M. (2010). Eficácia dos processos de linkagem na avaliação educacional em larga escala. *Estudos em Avaliação Educacional*, 21(45), 191-211.
- Souza, N. A. (2005). Avaliação de competências: o aperfeiçoamento profissional na área de enfermagem. *Estudos em Avaliação Educacional*, 16(32), 57-80.
- Tejada, A. J. R., & Meléndez, C. P. (2001). *Nuevos modelos para la medición de Actitudes*. Valencia: Promolibro.
- Valle, R. C. (2000). Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, 21, 7-91.
- Vendramini, C. M. M. (2005). Aplicação da Teoria de Resposta ao Item na avaliação educacional. Em R. Primi (Org.). *Temas em avaliação psicológica*, (pp. 229-254). São Paulo: Casa do Psicólogo.

Recebido em 04/07/2012  
 Reformulado em 12/02/2013  
 Aprovado em 10/04/2013

Nota das autoras:

Este artigo é parte da dissertação de mestrado da primeira autora orientada pela segunda e contou com o fomento da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes

Sobre as autoras:

**Fernanda Luzia Lopes** é psicóloga, doutoranda em Psicologia na Universidade São Francisco, com bolsa do Obeduc/Capes/Inep.

**Claudette Maria Medeiros Vendramini** é estatística. Doutora em Educação pela Universidade Estadual de Campinas, é professora associada doutora da graduação e Pós-Graduação em Psicologia da Universidade São Francisco.

Contato com as autoras:

Universidade São Francisco

Rua Alexandre Rodrigues Barbosa, 45 – Centro – CEP 13251-900 – Itatiba/SP, Brasil

E-mail: [claudette.vendramini@usf.edu.br](mailto:claudette.vendramini@usf.edu.br)