

Propriedades psicométricas das provas de pedagogia do ENADE via TRI

Fernanda Luzia Lopes

Claudette Maria Medeiros Vendramini

Resumo: Avaliação em larga escala no contexto educacional tem por objetivo investigar as habilidades e competências do estudante na realização de tarefas acadêmicas. Um dos testes acadêmicos responsável por avaliar o conhecimento específico em distintas áreas do conhecimento é o Exame Nacional de Desempenho dos Estudantes – ENADE, aplicado aos estudantes desde 2004. Desse modo, o presente artigo tem por objetivo avaliar as propriedades psicométricas da prova de Pedagogia do ENADE, aplicada no ano de 2005, por meio do modelo de Rasch da Teoria de Resposta ao Item. Para tanto, foi utilizado um banco de dados contendo informações acadêmicas de 49.497 estudantes do curso de Pedagogia que realizaram o ENADE em 2005. Como material utilizou-se as provas de Formação Geral e de Componente Específico, compostas por questões objetivas e dissertativas. Os resultados do estudo sugeriram itens com boas propriedades psicométricas, foi possível observar que os índices de dificuldade dos itens foram, em média, de dificuldade mediana. Quanto às pessoas, a maior parte evidenciou nível de habilidades compatíveis à dificuldade média dos itens (-0,14), indicando que a prova possui um nível de dificuldade semelhante ao nível de habilidade das pessoas. Ponderando a importância da avaliação educacional, estudos que possam averiguar a qualidade da avaliação do Ensino Superior são fundamentais considerando a relevância que o tema impõe para a construção de uma educação com melhores propriedades e reconhecimento no panorama sócio-político-econômico.

Palavras-chave: Avaliação em larga escala. Ensino superior. Métodos e medidas em psicologia.

Psychometric properties of ENADE's pedagogy test via TRI

Abstract: The large-scale assessment in the educational context aims at investigating the students' abilities and competences to do college tests. One of the academic tests responsible for assessing the specific knowledge in different areas is the National Students Performance Exam – ENADE given to students since 2004. This way, this article's objective is to assess the psychometric properties of the 2005 ENADE's Pedagogy test via the Rasch model of the Item Response Theory. To do so, we used a database containing academic information of the 49.497 Pedagogy course students who took the ENADE's exam in 2005. The material used for the analysis consisted of the General Formation and the Specific Component Exams, both had objective and essay questions. The results of this study suggest items with good psychometric properties. The mean of the items difficulty indexes showed an overall median difficulty level. Regarding the students, most of them presented an abilities level that was comparable to the mean items difficulty ($b=-0,14$), indicating that both the exams difficulty levels and the students ability had a similar level. Reflecting on the importance of educational assessment, studies investigating the quality of higher education assessments are fundamental, taking into account the relevance of the subject when it comes to building up quality education and recognition within the social-political-economic scenario.

Key words: Large-scale assessment. Higher education. Psychology methods and measures.

Introdução

O tema Educação vem sendo amplamente discutido nos últimos anos, considerando sua importância para a contemporaneidade, bem como para todos os aspectos que envolvem a sociedade. A educação é parte na tomada de decisões e de políticas públicas que se configuram em um fecundo campo de pesquisas, com possibilidades e compromissos de melhor compreender, discutir e privilegiar seu desenvolvimento. A primazia dessas discussões refere-se à qualidade da educação ofertada à população, dada a necessidade de um ensino substancialmente efetivo, haja vista seus efeitos nos mais diversos setores e camadas que compõem a organização sócio-político-econômica de um país.

Ao se tecer considerações acerca da qualidade do ensino vigente, a busca de instrumentos que viabilizem a avaliação criteriosa, pautados em ferramentas coerentes com a seriedade que o tema deve ser tratado, vem sendo o objetivo dos sistemas governamentais responsáveis pelo aprimoramento, expansão e acessibilidade da educação brasileira. Assim, para a avaliação do ensino superior, foco deste trabalho, o sistema em vigor é o SINAES (Sistema Nacional de Avaliação da Educação Superior), instituído em 2004, sendo o ENADE (Exame Nacional de Desempenho dos Estudantes) um instrumento componente desse sistema maior de avaliação, que busca avaliar os ingressantes e o ganho acadêmico de concluintes de cursos superiores. A partir dos resultados do ENADE e de outros indicadores educacionais, o SINAES visa avaliar cursos e Instituições de Ensino Superior (IES). Vale ressaltar que muitas mudanças têm ocorrido nos últimos anos quanto à forma de elaboração e de aplicação do ENADE, e que afastam o atual sistema de Avaliação da proposta inicial do SINAES. O ENADE, atualmente, não é mais aplicado em estudantes ingressantes e concluintes dos cursos de graduação, mas somente nos concluintes. A nota dos ingressantes foi substituída pela nota do ENEM (Exame Nacional do Ensino Médio), cuja finalidade não é a mesma do ENADE.

As provas do ENADE avaliam as habilidades e competências do estudante de graduação de acordo com sua área de ênfase. Um conceito de habilidade que se assemelha aos propósitos deste estudo é o apresentado por Brito et al. (2000), em que a habilidade pode ser compreendida como uma característica psicológica singular que trabalha em favorecimento aos domínios de uma determinada atividade, de forma rápida e fácil. Na situação acadêmica, a habilidade do estudante seria medida na atribuição de respostas corretas e/ou do conhecimento de determinados conteúdos apresentados, por exemplo, num teste de desempenho acadêmico.

Já o termo competência, de acordo com Brito et al. (2000), é de uso recente nas pesquisas educacionais e psicológicas, haja vista que as primeiras utilizações dessa terminologia datam a partir da década de 1970. Sua conceituação está ligada à ideia de desempenho eficiente na realização de tarefas que requerem habilidades cognitivas específicas. A relação entre competência e habilidade pode ser entendida como um processo em que, por exemplo, a competência para a realização de um teste de desempenho acadêmico depende de habilidades cognitivas coerentes com a demanda solicitada.

A habilidade, segundo Primi et al. (2001), pressupõe facilidade em manipular tipos de informações, e para que a habilidade se transforme em competência faz-se necessário um investimento em experiências anteriores de aprendizagem, caso contrário, não haverá competência. Desta maneira, alguns fatores estão associados à aquisição de competências, ao desenvolvimento de habilidades e à qualidade das experiências de aprendizagem vivenciadas pelo estudante ao longo de sua trajetória acadêmica (PRIMI et al., 2001).

Cabe ressaltar que, para o presente trabalho, será adotada a terminologia para habilidades acadêmicas e competências propostas por Brito (2008, p. 846), consonantes com o contexto avaliativo do ENADE, em que

Essa habilidade acadêmica é, então, a capacidade de um indivíduo realizar determinadas tarefas, solucionar determinados problemas, dominar com sucesso determinadas exigências do meio, obtendo e demonstrando domínio do conhecimento e de tarefas relativas a uma determinada atividade. Esse aspecto é importante porque não é possível observar uma habilidade em sua forma pura, pois, a habilidade se manifesta durante a execução de uma atividade. O que pode ser observado são manifestações dos componentes de uma determinada habilidade. Já a competência profissional é a capacidade de mobilizar, articular e colocar em ação conhecimentos, habilidades, atitudes e valores necessários para o desempenho eficiente e eficaz de atividades requeridas pela natureza do trabalho e do desenvolvimento tecnológico.

No âmbito da avaliação educacional em larga escala, Primi et al. (2001) destacam que as IES buscam o desenvolvimento das habilidades e das competências necessárias à condição de avaliação. Para Souza (2005), a formação de competências e o desenvolvimento de habilidades indicam um trabalho diferenciado das instituições de ensino preocupadas com as contribuições feitas à vida acadêmica de seus estudantes e, em consequência, seu desempenho acadêmico.

Soares, Ribeiro e Castro (2001) destacam que o desempenho acadêmico pode ser influenciado pelo ambiente cultural vivenciado pelo estudante, o envolvi-

mento da família no processo de aprendizagem, situação socioeconômica, bem como características individuais que podem incluir a motivação para o estudo, habilidades cognitivas e desempenho prévio em níveis de ensino anteriores. Os autores apontam para o fato de que as instituições de ensino devem levar esses fatores em consideração, visto que são agentes facilitadores na aquisição de conhecimento.

Além de considerar todas essas variáveis é importante que os instrumentos utilizados nessas avaliações sejam de qualidade, isto é, com propriedades psicométricas que garantam a validade e a confiabilidade de seus resultados. Dentre os modelos matemáticos e estatísticos que podem ser utilizados para verificar as propriedades psicométricas correspondentes a um exame educacional, pode-se destacar a Teoria de Resposta ao Item (TRI). A TRI vem sendo difundida por diversos autores a partir dos anos de 1950, como assinalam Pasquali e Primi (2007). Com o advento da tecnologia da informática, em meados dos anos de 1980, *softwares* apropriados foram criados para suprir as exigências das complexas operações matemáticas da TRI e, assim, aprimorar a viabilidade de seu uso.

Na TRI, a probabilidade de resposta correta depende do traço do sujeito, ou seja, o nível de habilidade do respondente (θ) e dos parâmetros dos itens (a_i , b_i e c_i). Assim é necessário estimar os valores destes parâmetros, com base nas respostas dos sujeitos aos itens. A esse processo é dado o nome de calibração ou parametrização, e é feito com o auxílio de programas específicos. O campo das pesquisas na área educacional que empregam os métodos recomendados pela TRI segue em grande efervescência, já que o advento de novas tecnologias promove sua utilização segura e confiável. Exemplos do uso da TRI na avaliação do Ensino Superior são propostos por Primi et al. (2009), Vendramini (2009) e Primi et al. (2010).

O estudo sugerido por Primi et al. (2009) teve por objetivo estabelecer critérios de desempenho com base na análise dos itens da prova de Psicologia do ENADE 2006, por meio do procedimento de referência ao item da TRI. Para tanto, os pesquisadores utilizaram um banco de dados contendo as respostas ao ENADE de 26.613 estudantes do curso de Psicologia. O instrumento utilizado foi a prova do ENADE, contendo 40 questões divididas em 10 do componente de Formação Geral e 30 questões que mediam o Conhecimento Específico. Cada item respondido corretamente recebeu 1 ponto e, para cada item discursivo, foi calculada uma nota decimal entre 0 e 1. O modelo de Rasch foi empregado para o cálculo de uma escala de habilidades e, por fim, transformada em uma nota de 0 a 100, de acordo com a porcentagem de acertos dos estudantes. Os resultados indicaram que as medidas de dificuldade dos itens e das habilidades dos

estudantes não apresentaram respostas inesperadas. Para estabelecer os pontos de corte das notas, foi utilizado o mapa de itens que possibilitou a distinção de quatro faixas de desempenho. A Faixa 1 e 2 caracterizam competências consideradas insuficientes aos concluintes do curso; a Faixa 3 considera o nível mínimo de competência esperado dos estudantes; já a Faixa 4, contém itens que exigem conhecimentos mais complexos da área, considerados adequados à atuação dos profissionais dessa área do conhecimento. As análises dos resultados permitiram observar que 72,9% dos estudantes concluintes do curso de Psicologia estão na Faixa 3, faixa considerada de compreensão mínima dos conteúdos abarcados pelo exame – aquém do que se espera de um profissional graduado nesse curso (PRIMI et al., 2009).

Na pesquisa desenvolvida por Vendramini (2009), a autora estuda a dimensionalidade do componente de Formação Geral que compõe o ENADE por meio da Análise Fatorial Completa baseada na TRI. A pesquisadora utilizou-se de um banco de dados contendo os resultados de 135.008 estudantes que participaram do ENADE em 2004 de diferentes áreas do conhecimento (Agronomia, Educação Física, Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Medicina, Medicina Veterinária, Nutrição, Odontologia, Serviço Social, Terapia Ocupacional e Zootecnia). Esse estudo também contemplou mais dois bancos de dados sendo um deles referentes às respostas de 268.448 universitários de áreas do conhecimento distintas que participaram do ENADE 2005 (Arquitetura e Urbanismo, Biologia, Ciências Sociais, Computação, Engenharias, Filosofia, Física, Geografia, História, Letras, Matemática, Pedagogia e Química). O outro banco de dados foi composto por 401.117 estudantes que realizaram o ENADE em 2006, de diferentes carreiras (Administração, Arquivologia, Bibliotecnomia, Biomedicina, Ciências Contábeis Econômicas, Comunicação Social, Design, Direito, Música, Normal Superior, Psicologia, Secretariado Executivo, Teatro e Turismo). Como instrumento foi utilizado o componente de Formação Geral das provas do ENADE 2004, 2005 e 2006, contendo 10 itens em cada uma dessas provas.

Os resultados encontrados por Vendramini (2009) indicaram que os estudantes do ano de 2005 obtiveram melhores resultados com relação aos estudantes dos outros anos; contudo, devido ao fato de serem provas diferentes em cada um dos anos, não se pode atribuir melhor desempenho aos alunos de 2005. No que se refere à consistência interna das provas, os coeficientes de Kuder-Richardson foram baixos (0,353, 0,393 e 0,255, respectivamente para os anos de 2004, 2005 e 2006), resultados já esperados, haja vista a abrangência da prova de Formação Geral que aborda diversos temas e requer diferentes

habilidades acadêmicas dos participantes. Já os resultados da análise fatorial *Promax* evidenciaram estrutura fatorial de cinco fatores não correlacionados entre si para a prova de 2004, explicando 45,5% da variância total dos dados. Para a prova de 2005 observou-se uma estrutura fatorial de quatro fatores que explica 37,9% da variância total. Para o ano de 2006 constatou-se uma estrutura máxima de dois fatores, com 18,1% de variância explicada pelos dois fatores. Assim, a prova de 2006 apresentou menor consistência e com poucos itens correlacionados entre si, o que impede a composição de uma estrutura fatorial adequada (VENDRAMINI, 2009).

Em outra pesquisa, Primi et al. (2010) buscaram detectar itens com funcionamento diferencial (DIF), ou seja, itens que apresentavam problemas em medir o desempenho de estudantes ingressantes e concluintes e, ainda, o desempenho dos estudantes de instituições públicas ou privadas de ensino. Para tanto, utilizaram-se de um banco de dados composto por 26.613 estudantes participantes do ENADE de Psicologia do ano de 2006. Assim, por meio da análise de Rasch e de regressão logística, os autores puderam perceber que, referente aos estudantes ingressantes e concluintes, cinco itens não atingiram o critério mínimo de DIF, sendo que três itens favoreceram os ingressantes e dois itens favoreceram os concluintes. Já para os estudantes de universidades públicas e privadas, sete itens atingiram o critério mínimo de DIF, em que 6 itens favoreceram estudantes de instituições públicas. Segundo os autores, um grupo menor de itens com DIF indicaram problemas em sua formulação, favorecendo os estudantes com menores níveis de habilidade. Contudo, a maior parte dos itens que apresentaram DIF também tinha por característica propriedades psicométricas adequadas. Dessa maneira, os autores puderam concluir a existência de um pequeno número de itens potencialmente capazes de produzir um viés à prova, haja vista que esses itens se referem mais aos sistemas de ensino do que aos momentos de formação do aluno, ingressantes ou concluintes.

Considerando as diferentes aplicações do modelo de Rasch nas pesquisas citadas, também para este estudo foi utilizado, para a análise dos dados, o modelo logístico de um parâmetro, ou seja, o modelo de Rasch. Por todo o exposto, pode-se dizer que, o modelo de Rasch é um exemplo prático de como uma medição pode ser aplicada aos dados empíricos, já que contempla propriedades estatísticas desejáveis a um instrumento de pesquisa de boa qualidade, permitindo uma aplicação prática na estimativa de parâmetros (PERLINE, WRIGHT, WAINER, 1979). Desse modo, o presente artigo tem por objetivo avaliar as propriedades psicométricas da prova de Pedagogia do ENADE, aplicada no ano de 2005, por meio do modelo de Rasch ou modelo de um parâmetro da 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 Ministério da Educação (MEC) para a instituição de Observatórios da Educação em IES, apoiados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior do MEC (CAPES/MEC). O Observatório da Educação, particularmente o da Universidade São Francisco, tem por objetivos prioritários buscar evidências de validade da estrutura interna do ENADE, compreender o sentido das dimensões encontradas empregando métodos da Teoria de Resposta ao Item (TRI) e, ainda, investigar as evidências de validade do exame por intermédio da verificação de suas 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 monitorado pelo MEC. Os bancos de dados são compostos por estudantes selecionados de acordo com os critérios e rigores do MEC, selecionados aleatoriamente por amostragem estratificada e contam com informações referentes ao curso, às IES, respostas ao Componente de Formação Geral, ao Componente Específico por curso e ao Questionário Socioeconômico.

Participantes

Participaram do ENADE, no ano de 2005, 49.497 estudantes de Pedagogia de todos os estados brasileiros e Distrito Federal, sendo que a maior concentração refere-se à região sudeste do país, divididos em 23.318 ingressantes (52,9%) e 26.179 concluintes (47,1%), com idade média de 30,05 anos ($DP = 8,88$). A maioria dos ingressantes é do gênero feminino (92,6%), assim como dos concluintes (92,0%). De acordo com a categoria administrativa das IES a maior parte (70,7%) dos participantes está matriculada em Instituições de Ensino Superior privadas.

Material

Serviu de material a prova do curso de Pedagogia do ENADE realizado em 2005. Cabe observar que a prova do ENADE de Pedagogia é composta por 40

questões no formato discursivo e objetivo de múltipla escolha, sendo que 10 são de avaliação de conhecimentos gerais – a prova de Formação Geral (FG); e, por 30 questões específicas da área, referentes ao Componente Específico (CE). Para este estudo foram utilizadas as questões objetivas, sendo 7 questões de FG e 26 questões do CE, ressalta-se que duas questões do CE (10 e 35) foram anuladas pela comissão do ENADE.

A prova de FG é comum a todos os cursos e tem por objetivo investigar a formação de um profissional ético e comprometido com a sociedade, bem como aferir o domínio de conhecimentos diversificados de habilidades acadêmicas e competências profissionais. Já o CE de cada área do conhecimento é elaborado com base nas Diretrizes Curriculares Nacionais (DCN), aprovadas pelo Conselho Nacional de Educação e no perfil profissional de cada carreira, levando em consideração os saberes fundamentais em cada área profissional (MEC/INEP, 2005).

Procedimento de análise dos dados

Os critérios utilizados na análise dos itens são embasados na Teoria de Resposta ao Item. Assim, o primeiro critério diz respeito à Correlação item total, em que são considerados ideais os itens com valores acima de 0,30; pode-se aceitar alguns itens a partir da correlação 0,20, se obedecerem aos demais critérios, assim como forem relevantes do ponto de vista do conteúdo avaliado. O segundo critério menciona as medidas de *Infit* e *Outfit*, o item para ser considerado adequado deve apresentar essas duas medidas com valores variando de 0,5 a 1,4, sendo considerado ideal quando não ultrapassa o valor de 1,2. Destaca-se que o indicador *Infit* é mais sensível para padrões de respostas inesperados em itens que, aproximadamente, têm uma dificuldade compatível com a habilidade da pessoa; e, o *Outfit*, mais sensível a observações inesperadas feitas por pessoas ao responderem itens muito fáceis ou muito difíceis para elas (LINACRE; WRIGHT, 1991).

O terceiro critério contempla a medida “*b*”, ou seja, o índice de dificuldade do item que pode ser muito fácil ($-4 < b \leq -2$), fácil ($-2 < b \leq -1$), mediano ($-1 < b \leq +1$), difícil ($+1 < b \leq +2$) e muito difícil ($+2 < b \leq +4$). E, por último, a questão das cargas dos resíduos completa o quarto critério. Vale destacar que os itens mais difíceis tendem a apresentar resíduos maiores, que os outros itens, e em alguns casos superiores a 0,30. Contudo, a medida ideal para a qualidade dos itens considera que a carga residual (*residual loading*) seja menor que 0,40 (LINACRE; WRIGHT, 1991).

Resultados e discussões

Antes da apresentação dos resultados obtidos por meio da TRI, faz-se necessária uma sucinta descrição de alguns aspectos que permearam a aplicação da prova, bem como, informações importantes que serviram de base para compor o banco de dados utilizado neste estudo. A prova do ENADE realizada no ano de 2005 com os estudantes do curso de Pedagogia teve duração de quatro horas e era composta por questões objetivas e dissertativas, divididas em um componente de Formação Geral e, outro, de Conhecimento Específico (MEC/INEP, 2005).

O componente de Formação Geral foi composto por 10 questões, três delas do tipo dissertativas, que objetivaram avaliar a compreensão de textos e figuras, a análise das informações apresentadas envolvendo as habilidades gerais, contextualizadas em situações contemporâneas dos estudantes selecionados para realizar o exame. Nessas questões buscou-se averiguar a capacidade dos estudantes em analisar, sintetizar, estabelecer relações, bem como, identificar contradições e construir hipóteses acerca das situações-problema apresentadas (MEC/INEP, 2005). Cabe ressaltar que, ao longo dos anos, o ENADE passou por modificações no que concerne aos objetivos pretendidos por suas provas. Essas modificações estão presentes nas Portarias do ENADE, por ano de aplicação, disponíveis no portal eletrônico do INEP (<http://www.inep.gov.br>).

Já no componente de Conhecimento Específico, constituído de 28 questões objetivas e duas dissertativas, elaborado pela Fundação Cesgranrio a partir das DCN dos cursos, teve-se por objetivos levantar indicadores da qualidade do ensino ofertado pelas IES além de buscar identificar os problemas e as necessidades provenientes da formação do Pedagogo. Esse componente da prova também procurou avaliar o desenvolvimento dos estudantes com relação a conhecimentos pedagógicos de formação geral, conhecimentos pedagógicos-didáticos e conhecimentos das áreas específicas. A Tabela 1 abarca os conteúdos solicitados aos estudantes nesse componente (MEC/INEP, 2005).

As estatísticas básicas dos resultados do ENADE são divulgadas em relatórios referentes a cada curso e disponibilizados na página eletrônica do INEP, mantida pelo MEC (<http://www.inep.gov.br>). Para ilustrar alguns desses resultados a Tabela 2 apresenta as estatísticas básicas da parte objetiva do componente de Formação Geral de Pedagogia de 2005, em que se é possível observar que os estudantes ingressantes apresentaram um desempenho semelhante ao desempenho dos estudantes concluintes, além de indicar uma variabilidade similar entre os dois grupos.

Tabela 1 - Conteúdos solicitados por áreas do conhecimento referentes à Pedagogia

Áreas	Conteúdos
Área da Formação Geral	<ul style="list-style-type: none"> - Filosofia da Educação; História da Educação; História da Pedagogia. - Sociologia da Educação. - Psicologia da Educação. - Teorias Pedagógicas. - Didática. - Avaliação do ensino e da aprendizagem. - Organização da Educação; Políticas Educacionais. - Pesquisa Educacional. - Tecnologia da Informação nas práticas educativas.
Específicas para docência	<ul style="list-style-type: none"> - Especificidades do trabalho docente. - Conteúdos e metodologias específicas (Educação Infantil; Alfabetização e Letramento; Língua Portuguesa e Literatura; Matemática; Ciências; Geografia; História; Linguagens Artístico-Culturais; Motricidade Humana; Educação de Jovens e Adultos). - Temas transversais.
Específicas para gestão escolar e outros espaços educativos	<ul style="list-style-type: none"> - Coordenação, elaboração e avaliação de projeto político pedagógico. - Implementação e avaliação do currículo. - Apoio pedagógico aos processos de ensino e de aprendizagem nos espaços educativos. - Organização de práticas de gestão. - Relação escola-comunidade e movimentos sociais. - Investigação, produção e difusão de conhecimentos pedagógicos e educacionais.

Fonte: MINISTÉRIO DA EDUCAÇÃO E CULTURA. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Relatório Técnico**. 2005. p. 10,11. Disponível em: <<http://www.INEP.gov.br/download/ENADE/2005/relatorios/>>. Acesso: 9 abr. 2009.

As questões de Formação Geral também passaram por tratamento estatístico em que se objetivou dividir as questões entre muito fáceis, fáceis, medianas, difíceis e muito difíceis. Desta forma, para que as questões fossem consideradas muito fáceis, o índice de acerto ao item deveria ser maior ou igual a 86%; para serem considerados fáceis, os itens deveriam estar na faixa de acerto de 85% a 61%; do mesmo modo, para serem considerados itens de dificuldade mediana, as questões deveriam obter um índice de acerto de 60% a 41%; já para

Tabela 2 - Estatísticas básicas das questões objetivas de Formação Geral e do Componente específico por grupo de estudantes segundo sua formação

Estatísticas	Valores das estatísticas por componente e grupo de estudantes					
	FG			CE		
	Total	Ingressante	Concluinte	Total	Ingressante	Concluinte
Média	63,6	62,5	64,6	45,7	41,5	49,4
Erro	00,1	00,1	00,1	00,1	00,1	00,1
Desvio-padrão	22,4	22,5	22,2	16,9	16,0	16,7
Nota mínima	00,0	00,0	00,0	00,0	00,0	00,0
Mediana	71,4	71,4	71,4	46,2	42,3	50,0
Nota máxima	100	100	100	96,2	96,2	92,3

Fonte: MINISTÉRIO DA EDUCAÇÃO E CULTURA. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Relatório Técnico. 2005. p. 35. Disponível em: <<http://www.INEP.gov.br/download/ENADE/2005/relatorios/>>. Acesso: 9 abr. 2009.

serem considerados itens difíceis, estes devem apresentar um índice de acerto de 16% a 40%; por fim, os itens avaliados como muito difíceis devem ter um índice de acerto menor ou igual a 15%. Na Tabela 3 são apresentados os itens correspondentes às questões objetivas do componente de Formação Geral, em que se é possível observar que a maior parte deles foi considerada fácil. Vale ressaltar que nas tabelas apresentadas pelo INEP, o índice de “dificuldade” é denominado como índice de “facilidade”, mas representa, nos dois casos, a porcentagem de acertos na questão (MEC/INEP, 2005).

Para o componente de Conhecimento Específico foram realizadas as mesmas análises em que se obtiveram as estatísticas apresentadas na Tabela 2. Conforme as informações contidas nessa tabela, os estudantes concluintes apresentaram um desempenho melhor ($M= 49,4$) do que o desempenho dos estudantes ingressantes ($M= 41,5$), muito embora a maior nota obtida neste componente fosse de estudantes ingressantes (96,2). A variabilidade das notas, como no caso do componente de Formação Geral, também foi similar nos dois grupos de estudantes.

Tabela 3 - Classificação das questões objetivas de Formação Geral e de Componente Específico por índice de facilidade

Facilidade (%)	Classificação	Questões	
		FG	CE
≥ 86	Muito fácil		17
61 a 85	Fácil	1, 2, 3, 4, 7	16, 21, 22, 32
41 a 60	Mediana	5	8, 12, 15, 19, 23, 25, 27, 30, 31, 33, 34
16 a 40	Difícil	6	9, 11, 14, 18, 20, 24, 26, 28, 29
≤ 15	Muito Difícil		13

Fonte: MINISTÉRIO DA EDUCAÇÃO E CULTURA. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Relatório Técnico. 2005. p. 43. Disponível em: <<http://www.INEP.gov.br/download/ENADE/2005/relatorios/>>. Acesso: 9 abr. 2009

O índice de facilidade da prova de Conhecimento Específico foi estabelecido segundo os mesmos critérios utilizados na determinação desses itens para o componente de Formação Geral. Estes índices e as respectivas questões estão apresentados na Tabela 3, indicando que a maior parte dos itens pode ser considerada de facilidade mediana, enquanto que, para essa parte da prova, um item foi avaliado como muito fácil (17) e um item avaliado como muito difícil (13). Ressalta-se que, para esta análise, os itens 10 e 25 não foram considerados, já que foram anulados pela comissão de avaliação do INEP (MEC/INEP, 2005).

Esses resultados corroboram as estimativas da qualidade dos cursos avaliados pelo ENADE. Sendo assim, por meio da medição das habilidades acadêmicas dos estudantes de Pedagogia que participaram do ENADE 2005, bem como de seus perfis profissionais, foi possível a categorização das IES's em conceitos. De acordo com o Relatório Síntese da área de Pedagogia (2005), esses conceitos são estabelecidos obtendo-se as notas finais dos estudantes ingressantes e concluintes do componente de Formação Geral (valendo 25% do conceito) e as notas finais da prova de Conhecimento Específico (valendo 75% do conceito, sendo que a nota dos ingressantes tem peso de 15% e dos

concluintes, peso de 60%). Já os cursos classificados como sem conceito (SC) são assim denominados por não contarem com ingressantes ou concluintes em seus cursos, o que torna inviável o cálculo das notas finais. Assim sendo, os conceitos se dividem em Conceito 1 (notas finais de 0,0 a 0,9), Conceito 2 (notas finais de 1,0 a 1,9), Conceito 3 (notas finais de 2,0 a 2,9), Conceito 4 (notas finais de 3,0 a 3,9), Conceito 5 (4,0 a 5,0).

De acordo com os critérios expostos para o cálculo dos conceitos dos cursos de Pedagogia no Brasil, avaliados pelo ENADE no ano de 2005, é possível afirmar, de acordo com a Tabela 4, que a maior parte dos cursos obtiveram o Conceito 3, o que equivale a 41,9% dos cursos participantes. A Tabela 4 também traz informações sobre as regiões do país e a respectiva distribuição do número de cursos por conceito, em que se pode observar que a região Sudeste apresentou mais cursos avaliados com o Conceito 3 (MEC/INEP, 2005).

Análises pela teoria de resposta ao item

Para a realização desta etapa do estudo, foi necessário selecionar, junto ao banco de dados fornecido pelo MEC, somente os estudantes do curso de Pedagogia. Esta tarefa foi possível com o auxílio do *software* estatístico *SPSS*, em sua versão 17.0. Depois de realizado esse procedimento, os dados foram tratados a fim de que pudessem ser manipulados pelo *software* *Winsteps*. Por meio deste programa, é possível a calibração dos parâmetros da TRI, de acordo com os aspectos do modelo de Rasch, em que foi relacionada a probabilidade de acerto ao item (nível de habilidade) com seu índice de dificuldade (LINA-CRE; WRIGHT, 1991).

Desta maneira, foi aplicada a TRI nas duas partes da prova do ENADE de Pedagogia 2005 em que, para o componente de Formação Geral obteve-se os resultados apresentados na Tabela 5. De acordo com essa tabela, a questão que demandou menor habilidade dos respondentes refere-se à Questão 4 ($b=-1,43$) que, dentre os critérios estabelecidos, pode ser considerado uma questão fácil. Já a questão que exigiu um nível maior de habilidade dos estudantes foi a Questão 6 ($b=1,92$) que pode ser considerada difícil. Esses resultados estão em consonância com os resultados divulgados pelo INEP em que a Questão 4 obteve uma porcentagem de acerto entre 61% e 85 %, avaliada como fácil e a Questão 6 que obteve uma porcentagem de acerto entre 16% e 40%, sendo o item mais difícil da prova (MEC/INEP, 2005).

Tabela 4 - Distribuição do número de cursos por conceito e regiões do Brasil no ENADE de Pedagogia de 2005

Conceito	Região											
	Brasil		Norte		Nordeste		Sudeste		Sul		Centro-Oeste	
	N	%	N	%	N	%	N	%	N	%	N	%
1	13	01,5	02	02,8	03	02,2	04	01,0	02	01,1	02	01,8
2	87	09,8	09	12,5	15	11,1	19	04,9	13	07,2	31	27,2
3	373	41,9	20	27,8	46	34,1	183	46,9	81	45,0	43	37,7
4	160	18,0	06	08,3	28	20,7	82	21,0	38	21,1	06	05,3
5	12	01,3	00	00,0	01	00,7	09	02,3	02	01,1	00	00,0
SC	246	27,6	35	48,6	42	31,1	93	23,8	44	24,4	32	28,1
Total	891	100,0	72	100,0	135	100,0	390	100,0	180	100,0	114	100,0

Fonte: MINISTÉRIO DA EDUCAÇÃO E CULTURA. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Relatório Técnico. 2005. p. 87. Disponível em: <<http://www.inep.gov.br/download/Enade/2005/relatorios/>>. Acesso: 9 abr. 2009.

Tabela 5 - Estatísticas dos itens em Formação Geral no ENADE de Pedagogia 2005 calibrados pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	Carga residual
01	30.128	44.496	-0,18	0,93	0,89	0,50	-0,41
02	33.920	44.624	-0,68	0,96	0,92	0,45	-0,16
03	29.500	44.618	-0,09	0,96	0,93	0,48	-0,15
04	38.238	44.620	-1,43	0,97	0,89	0,38	-0,16
05	21.820	44.592	0,78	1,04	1,06	0,46	0,97
06	12.194	44.550	1,92	1,09	1,23	0,43	-0,27
07	31.279	44.536	-0,32	1,04	1,05	0,41	-0,03
Média	28.154,1	44.576,6	0,00	1,00	0,99		
DP	7.975,4	46,2	1,00	0,05	0,12		

Com relação aos outros critérios estabelecidos para este estudo, os índices gerados por meio da TRI podem ser considerados aceitáveis, haja vista as medidas de *Infit* e *Outfit* que não foram inferiores a 0,5 e também não superaram o valor crítico de 1,40. Contudo, observou-se que a questão de número 6 atingiu na medida *Outifit* o valor de 1,23, valor que pode ser avaliado como marginalmente ideal, tendo por base que, o item para ser considerado ideal deve obter, nessa medida, o valor de 1,20. Da mesma forma, os índices da medida Correlação item total não foram inferiores ao valor crítico de 0,30. Já para a carga residual, em duas questões esses valores se apresentaram insuficientes, a saber, a questão de número 1 (-0,47) e a questão de número 5 (0,97).

Os mesmos procedimentos foram realizados para a obtenção da calibragem dos itens pelo modelo de Rasch do Componente Específico da Prova de Pedagogia do ENADE 2005. Os resultados estão apresentados na Tabela 6 em que se pode observar índices de dificuldade que indicam itens muito fáceis, fáceis, medianos, difíceis e muito difíceis. Com base nesses resultados e adotando os critérios pré-estabelecidos, o item que exigiu menor habilidade de seus respondentes foi o de número 17 ($b = -2,91$), considerado um item muito fácil. Já o item com maior índice de dificuldade foi o de número 13 ($b = 2,17$), representando um item muito difícil. Esses resultados também estão de acordo com os resultados divulgados pelo INEP, em que a Questão 17 obteve uma porcentagem de acertos maior ou igual a 86%, denominado como muito fácil, e a Questão 13 uma menor porcentagem de acertos ($\leq 15\%$), avaliada como muito difícil (MEC/INEP, 2005).

Conforme os dados referentes à Tabela 6 pode-se visualizar que todas as medidas de *Infit* e *Outfit* estão consonantes com os valores estabelecidos como suficientes que propõem a margem dessas medidas entre 0,5 e 1,4, indicando que não houve padrões de respostas inesperadas. Considerando o valor ideal de *Infit* e *Outfit* de 1,20, apenas um item, o de número 29, alcançou na medida de *Outfit* o valor de 1,22.

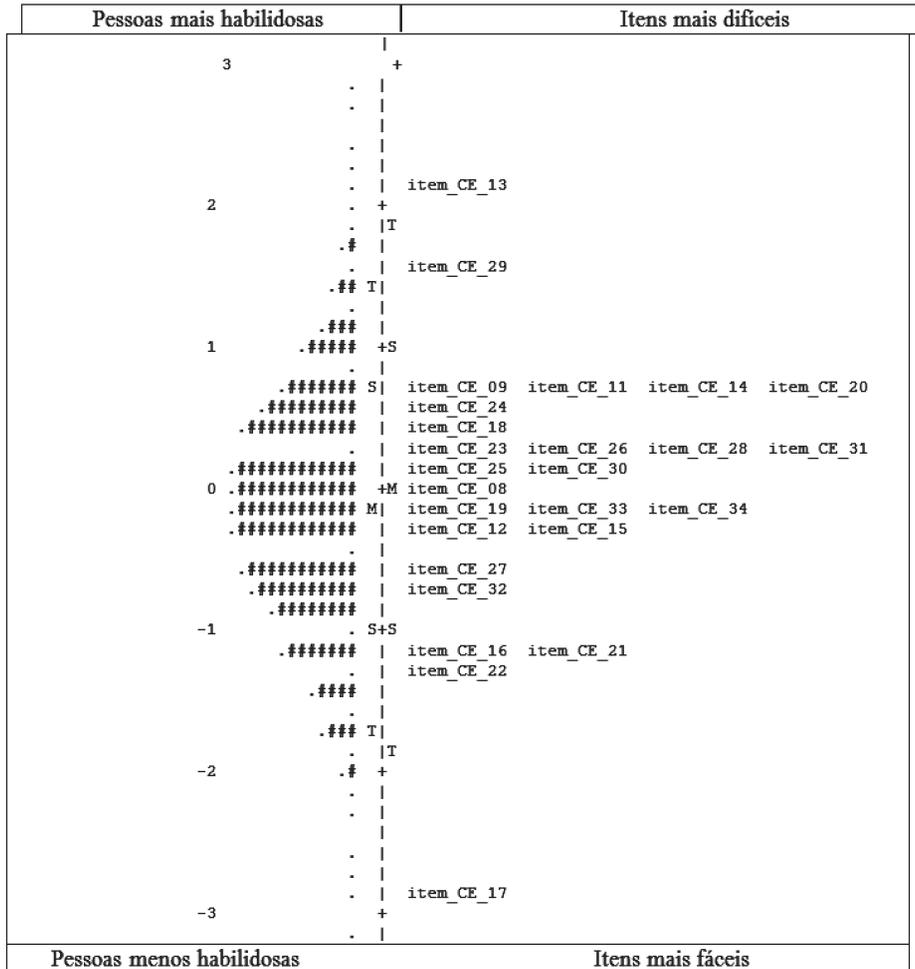
Quanto à associação entre os itens e a prova como um todo, os valores da medida entendida como Correlação item total, para os itens de número 11 ($r_{\text{item-total}} = 0,21$), 12 ($r_{\text{item-total}} = 0,28$), 13 ($r_{\text{item-total}} = 0,20$), 24 ($r_{\text{item-total}} = 0,24$), 25 ($r_{\text{item-total}} = 0,25$), 26 ($r_{\text{item-total}} = 0,23$), 29 ($r_{\text{item-total}} = 0,18$) e 31 ($r_{\text{item-total}} = 0,28$), tiveram suas correlações abaixo de 0,30, valor entendido como ideal, totalizando 30,77% dos itens. Cabe ressaltar que se pode aceitar como suficientes itens com medidas a partir de 0,20, se obedecerem aos demais critérios, bem como a relevância do conteúdo avaliado (PRIMI, 2006).

A Figura 1 representa o mapa de itens, uma das ferramentas da TRI, em que se pode observar a distribuição do nível de habilidade dos estudantes (à esquerda) e o nível de dificuldade dos itens (à direita). Assim, por meio do mapa de itens é possível observar que há uma concentração de estudantes entre os valores -1 a +1. O lado direito do gráfico corresponde à posição dos itens aplicados no ENADE de Pedagogia de 2005 em relação à sua dificuldade. A letra “M” de ambos os lados da figura, representa os valores médios de habilidade das pessoas (à esquerda) e de dificuldade dos itens (à direita), podendo ser observado que, no geral, o teste foi de dificuldade mediana para os estudantes avaliados. Os estudantes apresentaram uma média de habilidade próxima (um pouco abaixo) à média de dificuldade dos itens.

Tabela 6 - Estatísticas dos itens calibrados do Componente Específico do ENADE de Pedagogia 2005

Item	Score Bruto	N	Dificuldade	Infit	Outfit	Correlação Item-total	Carga residual
08	22.682	47.817	-0,01	1,02	1,02	0,32	-0,19
09	15.497	47.816	0,70	1,00	1,01	0,32	0,03
10	Anulado						
11	15.682	47.789	0,68	1,09	1,13	0,21	-0,33
12	24.809	47.656	-0,22	1,05	1,07	0,28	-0,09
13	5.327	47.838	2,17	1,01	1,13	0,20	0,01
14	15.502	47.806	0,70	0,94	0,93	0,39	0,32
15	25.457	47.889	-0,28	0,90	0,89	0,46	0,47
16	33.590	47.975	-1,09	0,88	0,82	0,48	0,31
17	44.349	47.962	-2,91	0,94	0,80	0,30	0,05
18	18.478	47.856	0,40	0,98	1,00	0,35	0,18
19	24.551	47.903	-0,19	0,94	0,93	0,42	0,26
20	15.638	47.919	0,69	1,02	1,03	0,30	-0,12
21	34.295	47.867	-1,18	1,00	0,97	0,33	-0,04
22	34.535	47.907	-1,22	0,94	0,92	0,39	0,09
23	20.180	47.905	0,23	1,01	1,02	0,33	0,16
24	16.345	47.774	0,61	1,07	1,10	0,24	-0,22
25	20.995	47.927	0,15	1,07	1,09	0,25	-0,34
26	19.810	47.844	0,26	1,08	1,12	0,23	-0,36
27	27.695	47.837	-0,52	0,92	0,90	0,44	0,20
28	19.979	47.847	0,25	0,96	0,95	0,39	0,18
29	8.943	47.870	1,51	1,05	1,22	0,18	0,01
30	21.142	47.842	0,13	0,98	0,98	0,36	0,09
31	20.262	47.859	0,22	1,05	1,05	0,28	-0,31
32	29.897	47.853	-0,71	0,97	0,95	0,38	0,15
33	24.407	47.704	-0,18	1,03	1,04	0,31	-0,11
34	24.477	47.820	-0,18	1,03	1,04	0,31	-0,24
35	Anulado						
Média	22.495,9	47.849,7	0,00	1,00	1,00		
DP	8.203,3	69,2	0,95	0,06	0,10		

Figura 1 - Comparação de habilidades dos estudantes e dificuldade dos itens



A Tabela 7 contempla as informações referentes aos parâmetros psicométricos dos itens da prova de Pedagogia do ENADE 2005 e de seus respondentes, sendo apresentado o ajuste médio das pessoas nas medidas de *infit* e *outfit*, a média de habilidade do grupo e os valores médios da dificuldade dos itens na prova e suas respectivas médias de ajuste.

Tabela 7 - Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas e itens

Referência	Estatísticas	Escore Bruto	Dificuldade	Infit	Outfit
Parâmetro dos itens	Média	22.495,0	0,00	1,00	1,00
	Desvio Padrão	8.203,3	0,95	0,06	0,10
	Máximo	44.349,0	2,17	1,09	1,22
	Mínimo	5.327,0	-2,91	0,88	0,80
Parâmetro das pessoas	Média	12,2	-0,14	1,00	1,00
	Desvio Padrão	4,0	0,80	0,15	0,31
	Máximo	25,0	3,58	1,79	9,90
	Mínimo	1,0	-3,70	0,44	0,11

Por esses dados, identifica-se que, em média, os itens apresentaram dificuldade mediana com indicadores que alcançaram os valores de dificuldade 2,17 (máximo) e -2,91 (mínimo). Também, no geral, os índices de ajuste de *infit* e *outfit* podem ser considerados aceitáveis. Quanto às pessoas, a maior parte evidenciou habilidades um pouco abaixo à da média dos itens (-0,14), o que pode indicar que a prova possui um nível de dificuldade semelhante à habilidade das pessoas. Com relação às habilidades dos estudantes, houve pessoas com habilidades muito elevadas, a mais alta alcançou o nível 3,58 de theta. Do mesmo modo, houve pessoas com baixa habilidade (theta = -3,70).

Considerações finais

A avaliação educacional tem demonstrado sua importância nos últimos anos no sentido de investigar a demanda e a qualidade do ensino no Brasil. Essa preocupação tem se estendido aos níveis de ensino Fundamental, Médio e Superior. No Ensino Superior a avaliação se deu inicialmente por meio das provas do Exame Nacional de Cursos/Provão e, mais recentemente, pelo ENADE. Com isso, estudos que possam averiguar a qualidade dessas provas são fundamentais considerando a relevância que o tema impõe para a construção da educação, tanto nas séries iniciais quanto no Ensino Superior, de melhor qualidade e reconhecimento no panorama sócio-político-econômico do país.

Assim, a TRI se apresenta como uma ferramenta eficaz para esse tipo de avaliação em larga escala, pois considera a habilidade dos sujeitos em suas respostas a cada item, promovendo bases mais sólidas e precisas para a discussão da questão educacional. Embora, a TRI como um conjunto de procedimentos de avaliação não seja totalmente estranha nos estudos referentes à Psicologia e à Educação, considerando seu uso pelo INEP desde a década de 1990, poucos e isolados estudos se apresentam à comunidade científica, o que impede uma discussão mais pormenorizada dos resultados obtidos. Cabe ressaltar que o primeiro congresso nacional específico em TRI foi realizado no final do ano de 2009, demonstrando a carência de mais estudos sobre o tema. Contudo, determinantes que visem uma melhor interpretação dos conteúdos propostos nesse estudo devem ser realizados a fim de uma melhor clareza dos problemas susceptíveis a área da avaliação educacional em larga escala.

Referências

BRITO, M. R. F. O SINAES e o ENADE: da concepção à implantação. **Avaliação**, Campinas; Sorocaba, v. 13, n. 3, p. 841-850, 2008.

BRITO, M. R.; MUNHOZ, A.; PRIMI, R.; GONÇALVEZ, M. H. et al. Exames Nacionais: uma análise do ENEM aplicado à Matemática. **Avaliação**, Campinas, v. 5, n. 4, p. 445-454, 2000.

LINACRE, J. M.; WRIGHT, B. D. **WINSTEPS**: Rasch-Model computer programs. Chicago: MESA Press, 1991.

MINISTÉRIO DA EDUCAÇÃO E CULTURA. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Relatório Técnico**. 2005. Disponível em: <<http://www.INEP.gov.br/download/ENADE/2005/relatorios/>>. Acesso: 9 abr. 2009.

PASQUALI, L.; PRIMI, R. Fundamentos da Teoria da Resposta ao Item – TRI. In: PASQUALI, L. (Org.). **Teoria de resposta ao item: TRI**. Brasília: Ed. UNB, 2007. p. 11-28.

PERLINE, R.; WRIGHT, B. D.; WAINER, H. The Rasch model as Additive Conjoint Measurement. **Applied Psychological Measurement**, Chicago, v. 3, n. 2, p. 237-255, 1979.

PRIMI, R. **A validade do ENADE para avaliação da qualidade dos cursos de instituições de ensino superior**. Projeto de Pesquisa. Itatiba: Universidade São Francisco, LabAPE, 2006.

PRIMI, R.; CARVALHO, L. F.; MIGUEL, F. K.; SILVA, M. C. R. Análise do funcionamento diferencial dos itens do Exame Nacional do Estudante (ENADE) de psicologia de 2006. **Psico-USF**, Itatiba, v. 15, n. 3, p. 379-393, 2010.

PRIMI, R.; NUNES, C. H. S. S.; SILVA, M. C. R.; BARTHOLOMEU, D.; CARVALHO, L. F.; MIGUEL, F. K.; VENDRAMINI, C. M. M. Aplicação da Teoria de Resposta ao Item na interpretação das notas do ENADE de Psicologia. **Revista de Educação ANEC**, Brasília, v. 38, n. 150, p. 115-124, 2009.

PRIMI, R.; SANTOS, A. A. A.; VENDRAMINI, C. M. M. et al. Competências e habilidades cognitivas: diferentes definições do mesmo construto. **Psicologia: Teoria e Pesquisa**, Brasília, v. 17, n. 2, p. 151-159, 2001.

SOUZA, N. A. Avaliação de competências: o aperfeiçoamento profissional na área de enfermagem. **Estudos em Avaliação Educacional**, São Paulo, v. 16, n. 32, p. 57-80, 2005.

SOARES, J. F.; RIBEIRO, L. M.; CASTRO, C. M. Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de direito, administração e engenharia civil. **Dados**, Rio de Janeiro, v. 44, n. 2, p. 363-396, 2001.

VENDRAMINI, C. M. M. Análise Fatorial baseada na TRI do componente de Formação Geral do ENADE. **Revista de Educação ANEC**, Brasília, v. 38, n. 150, p. 153-163, 2009.

Fernanda Luzia Lopes – Universidade São Francisco
Itatiba | SP | Brasil. Contato: fernandallopes@hotmail.com

Claudette Maria Medeiros Vendramini – Universidade São Francisco
Itatiba | SP | Brasil. Contato: claudette.vendramini@usf.edu.br

Artigo recebido em 19 de outubro de 2012
e aprovado em 11 de abril de 2013.

