

FERNANDA LUZIA LOPES



EQUALIZAÇÃO DE PROVAS DE HABILIDADES ACADÊMICAS:

ESTUDOS A PARTIR DO ENADE VIA TRI

ITATIBA

2011

FERNANDA LUZIA LOPES

EQUALIZAÇÃO DE PROVAS DE HABILIDADES ACADÊMICAS:

ESTUDOS A PARTIR DO ENADE VIA TRI

Dissertação apresentada ao Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco para obtenção do título de Mestre.

ORIENTADOR(A): CLAUDETTE MARIA MEDEIROS VENDRAMINI

ITATIBA

2011

378.1.001 F852e	Lopes, Fernanda Luzia. Equalização de provas de habilidades acadêmicas: estudos a partir do ENADE via TRI. / Fernanda Luzia Lopes. -- Itatiba, 2011. 131 p. Dissertação (mestrado) – Programa de Pós- Graduação <i>Stricto Sensu</i> em Psicologia da Universidade São Francisco. Orientação de: Claudette Maria Medeiros Vendramini. 1. Avaliação em larga escala. 2. Métodos estatísticos. 3. Modelo de Rasch. 4. Ensino superior. 5. Psicometria. I. Vendramini, Claudette Maria Medeiros. II. Título.
--------------------	--

Ficha catalográfica elaborada pelas bibliotecárias do Setor de
Processamento Técnico da Universidade São Francisco.



UNIVERSIDADE SÃO FRANCISCO
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU
EM PSICOLOGIA

Fernanda Luzia Lopes, defendeu a dissertação "Equalização de provas de habilidades acadêmicas: estudos a partir do ENADE via TRI" aprovada pelo Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco em 1º de agosto de 2011 pela Banca Examinadora constituída por:

Prof. Dra. Claudette Maria Medeiros Vendramini (Presidente)
Universidade São Francisco

Prof. Dr. Ricardo Primi
Universidade São Francisco

Prof. Dra. Marcia Regina Ferreira de Brito Dias
Universidade Estadual de Campinas

*À memória de
Escolástica Aparecida Andrade Bazeto
e José Bazeto
de onde estiverem...*

*(...)Deus nos dá pessoas e coisas, para aprendermos a alegria..
Depois, retoma coisas e pessoas para ver se já somos capazes
da alegria sozinhos... Essa a alegria que Ele quer.*

*(...)Só se pode viver perto do outro, e conhecer outra pessoa sem perigo de
ódio, se a gente tem amor. Qualquer amor já é um pouquinho de saúde, um
descanso na loucura.
(João Guimarães Rosa)*

AGRADECIMENTOS

Ao meu grande amor, Éder, que a cada dia me mostra o quanto a vida é melhor se a ‘gente andar mais à toa’, se deixarmos o amor e a simplicidade serem mais presentes no presente. Obrigada pelo (re)encontro, pela compreensão, pelo carinho, pela ‘leveza’, pelas ‘brincas’ e por toda Poesia que nos rodeia. Mon cher!

À minha mãe, Nazi, e ao meu irmão, Juju. Por sentirem orgulho do que faço, mesmo sem saber, exatamente, o que faço. Amo-os, mesmo sem o uso de palavras.

À minha ‘mãe acadêmica’, Prof^a Dr^a Claudette, que desde o meu primeiro ano de graduação não poupou esforços em dividir seu conhecimento, sempre de forma carinhosa e paciente. Obrigada por tudo!

Ao Prof^o Dr^o Ricardo Primi, por abrir as portas à pesquisa e pelas grandes contribuições a este trabalho.

Aos professores da Pós-Graduação, em especial aos professores Fermino Sisto, Maria Cristina Joly, Makilim Baptista, Cláudio Capitão, pelas sugestões, pelo carinho e pelo incentivo.

À Professora Dr^a Márcia Brito pelas valiosas sugestões que possibilitaram importantes mudanças neste trabalho.

Aos colegas do LabAPE, Marjorie, Daniel, Sanyo, Rodnei, Andréia, Fernando, Lucas e Adriano.

Aos queridos amigos que sempre estiveram próximos e que puderam compreender minha ausência e compartilhar minhas ansiedades.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, por tornar possível a realização deste trabalho.

RESUMO

Lopes, F. L. (2011). *Equalização de provas de habilidades acadêmicas: Estudos a partir do ENADE via TRI*. Dissertação de Mestrado. Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba.

Avaliação em larga escala no contexto educacional tem por objetivo investigar as habilidades e competências do estudante na realização de tarefas acadêmicas. A habilidade pode ser compreendida como uma característica psicológica que favorece o domínio de atividades específicas de forma rápida e fácil. Nesse contexto, as habilidades exigidas do estudante seriam medidas por meio do sucesso nas respostas assertivas de um teste acadêmico. 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. O presente estudo teve por objetivo equalizar provas via Teoria de Resposta ao Item – TRI a partir desse exame, para o curso de Pedagogia. 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 e um banco de dados contendo informações acadêmicas de 260 estudantes, divididos em Campus 1 e 2, de uma Instituição de Ensino Superior – IES particular paulista, que realizaram uma prova de Pedagogia de um Programa de Avaliação Continuada – PAC. Foram realizados três estudos, o primeiro contemplou a análise dos itens do ENADE e a seleção dos itens para a prova PAC, o segundo estudo privilegiou a análise da prova PAC e o terceiro estudo promoveu a comparação entre desempenhos dos estudantes do ENADE e do PAC. Os resultados do primeiro estudo sugeriram itens com boas propriedades psicométricas dos quais 16 deles foram selecionados como itens âncoras para comporem a prova. No segundo estudo foi possível observar que os índices de dificuldade dos itens foram, em média, de dificuldade mediana (-0,13). Quanto às pessoas, a maior parte evidenciou nível de habilidades compatíveis à dificuldade média dos itens (-0,16), indicando que a prova possui um nível de dificuldade semelhante ao nível de habilidade das pessoas. O terceiro estudo apontou que o desempenho dos estudantes concluintes da IES na prova do PAC obteve notas médias percentílicas semelhantes ao desempenho dos estudantes brasileiros de Pedagogia com Conceito 2 que realizaram o ENADE em 2005. Os estudantes concluintes da IES particular que realizaram o PAC alcançaram notas médias semelhantes às notas médias dos estudantes brasileiros concluintes com Conceito 4 que realizaram o ENADE. 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 atual.

Palavras-chave: Avaliação em Larga Escala, Métodos Estatísticos, Modelo de Rasch, Ensino Superior, Psicometria.

ABSTRACT

Lopes, F. L. (2011). *Equalization of tests of academic skills: Studies from the ENADE via TRI*. Master's degree dissertation. Postgraduate program *Stricto Sensu* in Psychology, Sao Francisco University, Itatiba.

Large scale evaluation in the educational context aims to investigate the skills and competences of the student performing academic tasks. Skill can be understood as a psychological characteristic which favours the domain of specific activities, making one perform them quickly and easily. In that context, the demanded skills for the student would be measured by the success obtained with assertive answers of an academic test.

One of the academic tests responsible for evaluating the specific knowledge in distinct areas of the knowledge is the Students' Performance National Exam, ENADE (abbreviation in Portuguese), taken by the students since 2004. This study aims to equalize tests via Item Response Theory – TRI (abbreviation in Portuguese) from that exam, in the Pedagogy. Therefore, a database storing academic information of 49.497 Pedagogy students who took the ENADE in 2005 and another database storing academic information of 260 students, divided in Campus 1 and 2, of a Higher Studies Institution – private IES (abbreviation in Portuguese) of the state of Sao Paulo, who took a Pedagogy test in a Continued Evaluation Program –PAC (abbreviation in Portuguese) were used. Three studies were done. The first comprehended the analysis of the items in the ENADE and the selection of them to the PAC test, the second focused on the analysis of the PAC test and the third study promoted the comparison between the performance of the students in the ENADE and PAC. The results of the first study suggested items with good psychometric properties and 16 of those were selected as linking items to be part of the test. In the second study it was possible to observe that the index of the difficulty in the items were of, in average, medium difficulty (-0,13). Regarding people, the biggest part showed compatible skills to the average difficulty in the items (-0,16), indicating that the test contains a level of difficulty similar to the people's skills. The third study focused on the performance of the students in the PAC test who were graduating in the IES. They have reached percentile average scores similar to the performance of the Brazilian students of Pedagogy with Concept 2 who took the ENADE in 2005. The students who were graduating the private IES who took the PAC have reached average scores similar to the average of the Brazilian students graduating with Concept 4 who took the ENADE. Considering the importance of the educational evaluation, studies that are able to check the quality of the evaluation in the Higher Studies level are necessary regarding the relevance that the theme imposes to the building of an education with better properties and to obtain award in the current socio-politic-economic scene.

Key-words: Large Scale Evaluation, Statistic Methods, Rasch Model, Higher Studies, Psychometric.

SUMÁRIO

LISTA DE FIGURAS	xii
LISTA DE TABELAS	xv
APRESENTAÇÃO	1
CAPÍTULO 1 - HABILIDADES E COMPETÊNCIAS.....	5
CAPÍTULO 2 - AVALIAÇÃO DO ENSINO SUPERIOR	18
HISTÓRICO DA AUTOAVALIAÇÃO DE UMA UNIVERSIDADE PARTICULAR DO INTERIOR DE SÃO PAULO	30
CAPÍTULO 3 - A TEORIA DE RESPOSTA AO ITEM.....	34
PESQUISAS COMO EMPREGO DA TRI NO BRASIL	43
EQUALIZAÇÃO E TRI	49
OBJETIVOS.....	53
OBJETIVO GERAL	53
OBJETIVOS ESPECÍFICOS	53
CAPÍTULO 4 - RESULTADOS E DISCUSSÃO	54
4.1 <i>ESTUDO 1</i> – ANÁLISE E SELEÇÃO DOS ITENS DO ENADE 2005 DE PEDAGOGIA	54
4.1.1 Método	54
4.1.2 Resultados e Discussão.....	59
4.2. <i>ESTUDO 2</i> – ANÁLISE DE UMA PROVA PAC DE PEDAGOGIA EQUALIZADA AO ENADE 2005	91
4.2.1 Método	91
4.2.2 Resultados e Discussão.....	93

4.3 <i>ESTUDO 3</i> – ANÁLISE E INTERPRETAÇÃO DOS RESULTADOS DO PAC COMPARADOS AO ENADE EM UMA MESMA ESCALA	104
4.3.1 Método	104
4.3.2 Resultados e Discussão	105
CAPÍTULO 5 - CONCLUSÃO E CONSIDERAÇÕES FINAIS	115
REFERÊNCIAS	118

LISTA DE FIGURAS

Figura 1 – Estrutura hierárquica das habilidades cognitivas de Cattell-Horn-Carroll.....	07
Figura 2 – Modelo logístico de três parâmetros.....	36
Figura 3 – Comparação de habilidades dos estudantes e dificuldade dos itens.....	07
Figura 4 – Questão de número 1 do PAC e correspondente à questão 2 do componente de Formação Geral do ENADE 2005.....	72
Figura 5 – Questão de número 2 do PAC e correspondente à questão 3 do componente de Formação Geral do ENADE 2005.....	73
Figura 6 – Questão de número 3 do PAC e correspondente à questão 4 do componente de Formação Geral do ENADE 2005.....	74
Figura 7 – Questão de número 4 do PAC e correspondente à questão 5 do componente de Formação Geral do ENADE 2005.....	75
Figura 8 – Questão de número 5 do PAC e correspondente à questão 7 do componente de Formação Geral do ENADE 2005.....	77
Figura 9 – Questão Correspondente ao item 9 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	79
Figura 10 – Questão Correspondente ao item 14 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	80

Figura 11 – Questão Correspondente ao item 15 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	81
Figura 12 – Questão Correspondente ao item 16 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	82
Figura 13 – Questão Correspondente ao item 17 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	83
Figura 14 – Questão Correspondente ao item 18 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	84
Figura 15 – Questão Correspondente ao item 20 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	85
Figura 16 – Questão Correspondente ao item 21 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	86
Figura 17 – Questão Correspondente ao item 23 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	87
Figura 18 – Questão Correspondente ao item 28 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	88
Figura 19 – Questão Correspondente ao item 34 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.....	89
Figura 20 – Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 03 da prova de Formação Geral do PAC.....	94

Figura 21 – Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 10 da prova de Formação Geral do PAC.....	95
Figura 22 – Comparação de habilidades dos estudantes e dificuldade dos itens.....	99
Figura 23 – Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 36 da prova de Componente Específico do PAC.....	100
Figura 24 – Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 32 da prova de Componente Específico do PAC.....	101
Figura 25 – Box-plot dos percentis de nota média no ENADE de ingresantes e concluintes por conceito do curso.....	107
Figura 28 – Box-plot dos percentis de nota média no PAC por ano e Campus.....	108

LISTA DE TABELAS

Tabela 1 – Descrição dos dez fatores amplos do modelo CHC e habilidades relacionadas.....	11
Tabela 2 – Distribuição dos estudantes e cursos submetidos ao ENADE em 2005.....	24
Tabela 3 – Diferenças entre a Teoria Clássica e a Teoria de Resposta ao Item.....	39
Tabela 4 – Distribuição de estudantes no ENADE 2005 de Pedagogia por categoria administrativa.....	55
Tabela 5 – Níveis de dificuldade dos itens quando estimados por TRI.....	57
Tabela 6 – Conteúdos solicitados por áreas do conhecimento referentes à Pedagogia.....	59
Tabela 7 – Estatísticas básicas das questões objetivas de Formação Geral por grupo de estudantes.....	60
Tabela 8 – Classificação das questões objetivas de Formação Geral por índice de facilidade.....	61
Tabela 9 – Estatísticas básicas das questões objetivas de Conhecimento Específico por grupo de estudantes.....	62
Tabela 10 – Classificação das questões objetivas do Componente Específico por índice de facilidade.....	63
Tabela 11 – Distribuição dos conceitos.....	64

Tabela 12 – Distribuição das regiões do Brasil por conceito do curso no ENADE de Pedagogia de 2005.....	64
Tabela 13 – Estatísticas dos itens em Formação Geral no ENADE de Pedagogia 2005 calibradas pelo modelo de Rasch.....	66
Tabela 14 – Estatísticas dos itens no Componente Específico do ENADE de Pedagogia 2005 calibradas pelo modelo de Rasch.....	68
Tabela 15 – Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas e itens.....	71
Tabela 16 – Critérios para a diferenciação da situação acadêmica.....	91
Tabela 17 – Distribuição dos estudantes pela situação acadêmica.....	91
Tabela 18 – Estatísticas dos itens em Formação Geral no PAC de Pedagogia calibradas pelo modelo de Rasch.....	93
Tabela 19 – Estatísticas dos itens no Componente Específico do PAC de Pedagogia calibradas pelo modelo de Rasch.....	97
Tabela 20 – Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas e itens.....	102
Tabela 21 – Distribuição de estudantes de Pedagogia por conceito no ENADE 2005.....	105
Tabela 22 – Distribuição de estudantes de Pedagogia que fizeram o PAC por Campus e situação acadêmica.....	106

APRESENTAÇÃO

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 fértil campo de pesquisas e de possibilidades, com o compromisso de 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.

Contudo, a preocupação com a qualidade do ensino ofertado não é somente privilégio do governo federal. Cada vez mais as instituições de ensino superior, motivadas

pelo SINAES, 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 do(s) instrumento(s) adotado(s). Diante disso, é nesse contexto que a Psicologia e os métodos psicométricos contribuem com a área Educacional, tornando possível a 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 Continuada (PAC) em uma universidade particular do interior paulista. O PAC é um sistema de avaliação institucional que conta com provas construídas de acordo com a matriz do ENADE e, ainda, com itens comuns à prova do ENADE, que possibilitam a comparação do desempenho acadêmico dos estudantes de sua instituição em relação aos estudantes brasileiros. Desde a instituição desse programa, as provas PAC vêm sendo assessoradas pelo Laboratório de Avaliação Psicológica e Educacional (LabAPE), vinculado ao Programa de Pós-Graduação *Stricto Sensu* da Universidade São Francisco e seus resultados são divulgados, anualmente, por meio de Relatórios Técnicos Institucionais.

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, demonstrando não somente

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.

O presente trabalho teve por objetivo geral equalizar provas de habilidades acadêmicas, a partir do ENADE via TRI. Os objetivos específicos contemplam verificar as propriedades psicométricas dos itens que compuseram a prova de Pedagogia do ENADE 2005; descrever o processo de elaboração de uma prova PAC composta por itens que apresentem as melhores propriedades psicométricas e itens indicados por professores especialistas na área de Pedagogia; análise e interpretação do mapa de itens obtidos pela TRI; e, verificar as possíveis diferenças no desempenho acadêmico dos estudantes que fizeram uma prova PAC e os que fizeram o ENADE 2005.

Para tanto, este estudo divide-se em seis capítulos. No Capítulo 1 são apresentadas as principais teorias das habilidades cognitivas em que se podem observar o percurso teórico, desde as primeiras pesquisas, e os conceitos de Spearman, seguido das ideias de Thurstone, Cattell, Horn e Carrol, que possibilitaram o modelo CHC da inteligência e, ainda, o acréscimo de mais seis habilidades cognitivas, dentre elas a *Gkn*, referente ao conhecimento específico, importante na realização de testes acadêmicos. Nesse Capítulo, também, são contemplados os conceitos de habilidades acadêmicas e competências, construtos considerados inerentes a avaliações de desempenho acadêmico, sejam estas em pequena ou em larga escala.

A avaliação do Ensino Superior é tema do Capítulo 2. Nesta seção são discutidos os percursos do governo federal na adoção de metodologias e implementação de sistemas de avaliação do ensino superior como é o caso do SINAES, atual modelo de avaliação, vigente

desde o ano de 2004. Também são consideradas neste capítulo discussões sobre os critérios e cuidados que devem ser monitorados quando se pretende uma avaliação abrangente e com rigores científicos na análise dos dados obtidos, haja vista a importância desses resultados para o contexto educacional.

No Capítulo 3 é exposta a Teoria de Resposta ao Item, seu histórico, suas principais diferenciações e vantagens sobre a Teoria Clássica dos Testes, seus pressupostos e utilizações. Dessa maneira são tratados tópicos que mencionam a importância da unidimensionalidade para a TRI; os modelos de um, dois e três parâmetros; pesquisas na área educacional que empregam a TRI para tratar seus resultados; e, ainda, a ferramenta da equalização de provas, foco desta pesquisa. Seguindo esse capítulo encontra-se o objetivo geral e os objetivos específicos propostos para este estudo.

A Metodologia utilizada nesta pesquisa, bem como os resultados obtidos em cada um desses estudos são contemplados no Capítulo de número 4, dividida em três estudos que melhor pudessem atingir os objetivos indicados. É também nesse capítulo em que são descritos os procedimentos e comentados os índices gerados com a aplicação da TRI. Por fim, o quinto Capítulo trata das considerações finais.

CAPÍTULO 1

HABILIDADES E COMPETÊNCIAS

Avaliações em larga escala no contexto educacional têm por objetivo investigar as habilidades acadêmicas que englobam os conceitos de desempenho e competência - dimensões humanas que são alvo de constantes pesquisas na área da Psicologia. Assim sendo, os sistemas de avaliação implementados pelo governo brasileiro tais como o Sistema Nacional de Avaliação do Ensino Básico – SAEB, o Exame Nacional do Ensino Médio – ENEM e o Exame Nacional do Desempenho dos Estudantes – ENADE, definem um conjunto de habilidades e competências essenciais aos diferentes estratos da educação (Primi & cols., 2001).

O conceito de habilidade apresentado por Carroll (1993) está associado à performance ou algum tipo de potencial de execução. O autor definiu a habilidade como um traço latente ou características de indivíduos que se manifestam diferentemente em uma classe de tarefas que variam em suas exigências (Carroll, 1992). Para ele, organizar uma habilidade ou uma medida de capacidade inclui a averiguação do tipo de tarefa e que atributos dessa tarefa devem ser envolvidos nessa medição. Outra questão a ser considerada menciona em determinar se a medida apresenta características desejáveis como, por exemplo, alta homogeneidade e confiabilidade, de acordo com os requisitos do modelo psicométrico utilizado.

Com base na realização de uma meta-análise dos artigos e trabalhos acerca das habilidades cognitivas, Carroll (1992) identificou 461 conjuntos de dados, pertinentes à temática, que o levaram a concluir que a estrutura das habilidades pode ser melhor descrita

por meio de um modelo de três camadas, ou estratos. Nesse estudo o autor utilizou-se de métodos da análise fatorial exploratória cujos resultados culminaram na Teoria dos Três Estratos, em que as habilidades mentais, ou inteligências, são disponibilizadas em três camadas organizadas hierarquicamente (Carroll, 1992; Carroll, 1993).

Na teoria proposta por Carroll (1992; 1993), o Estrato III é composto pelo fator *g*, isto é, o fator mais abrangente. O segundo Estrato (II), no qual se localizam as habilidades gerais, é composto por dez fatores gerais que influenciam uma gama de capacidades, tais como a de leitura e escrita, processamento visual, memória de curto prazo, inteligências fluída e cristalizada, entre outros. Já o Fator I é combinado por um conjunto de habilidades específicas, tais como o conhecimento léxico, a habilidade fonológica, a capacidade de visualização e vários tipos de habilidades referentes ao tempo de reação.

Na Figura 1 é apresentada, de forma esquemática, a estrutura hierárquica das habilidades cognitivas sugeridas por Cattell-Horn-Carroll em que, no Estrato III, encontra-se o fator geral da inteligência (*G*). O Estrato II é composto por dez fatores ligados às áreas do funcionamento cognitivo, a Inteligência Fluída (*Gf*), a Inteligência Cristalizada (*Gc*), o Conhecimento Quantitativo (*Gq*), a Leitura e Escrita (*Grw*), a Memória a Curto Prazo (*Gsm*), o Processamento Visual (*Gv*), o Processamento Auditivo (*Ga*), o Armazenamento e Recuperação Associativa a Longo Prazo (*Glr*), a Velocidade de Processamento Cognitivo (*Gs*) e o Tempo/Velocidade de Decisão/Reação (*Gt*). Já no Estrato I, encontram-se cerca de 70 fatores específicos que se relacionam ao conjunto de tarefas normalmente avaliadas por testes em geral (Carroll, 1993; McGrew, 2004; Schelini, 2006).

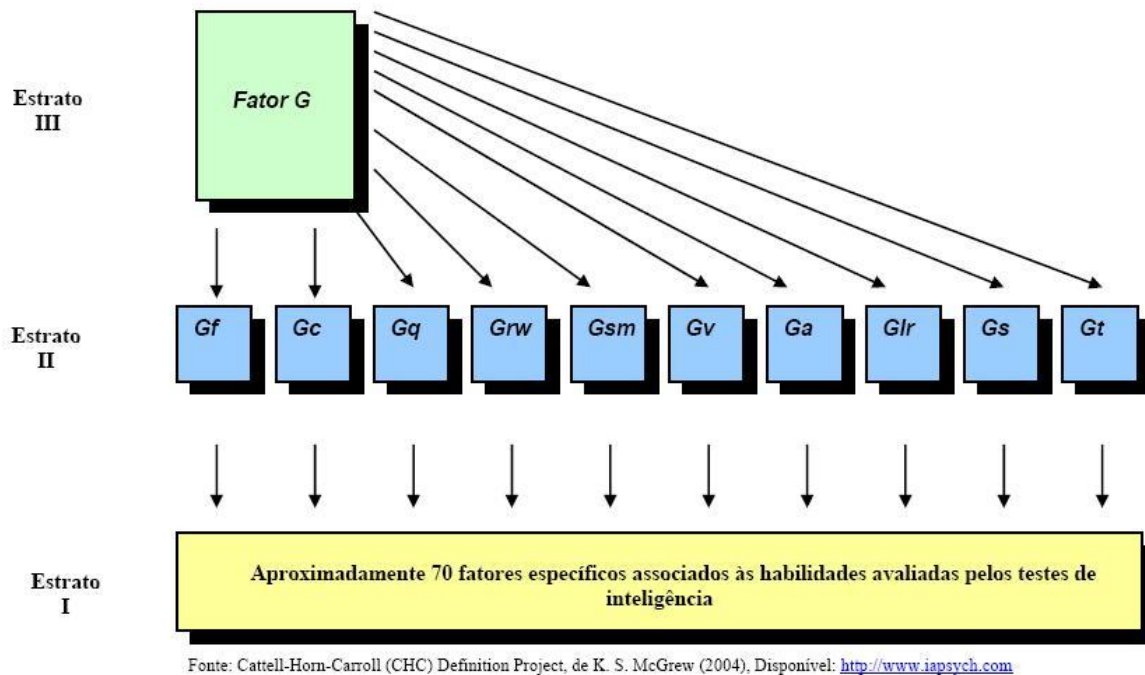


Figura 1. Estrutura hierárquica das habilidades cognitivas de Cattell-Horn-Carroll

A constituição de uma teoria das habilidades cognitivas se torna relevante para as questões da medida, em que a validade de testes cognitivos que têm por objetivo prever o sucesso escolar ou profissional, pode ser atribuída a um fator geral da inteligência (Carroll, 1992). Gottfredson (1986, citado por Carroll, 1992) já considerava que o fator geral da inteligência é a variável mais importante na previsão de sucesso educacional e profissional, sendo o sucesso e o desempenho educacional focos deste estudo.

As pesquisas de Carroll, segundo McGrew (2008), representam a primeira tentativa de uma taxonomia, com base empírica, de elementos das habilidades cognitivas organizadas estruturalmente. Para o autor, o modelo CHC não deve ser visto como um fim último da investigação psicométrica sobre as estruturas das capacidades humanas, mas sim ser um ponto de partida para novos estudos e investigações.

O modelo proposto por Carroll não foi a primeira tentativa de estabelecer uma teoria acerca das habilidades cognitivas. Muito antes, estudiosos como Spearman e Thurstone, desenvolveram métodos que procuravam investigar a estrutura e a definição das capacidades mentais, além dos fatores que respondiam pelo processo do pensamento, com base em modelos fatoriais da inteligência (Almeida, 1994; Ribeiro, 1998). Tais modelos eram extraídos por um conjunto de técnicas estatísticas, a análise fatorial, que permite a identificação dos fatores advindos das correlações entre variáveis observáveis (Hair, Anderson, Tetham & Black, 2005).

Cabe destacar que, ainda na primeira metade do século XX, as correntes do pensamento predominantes do conceito inteligência recaiam sobre os postulados de Spearman e Thurstone. Por um lado, Spearman considerava a existência de um fator *g* como princípio de uma energia mental, inata e comum aos indivíduos, que regia as tarefas cognitivas e era um fator determinante à aprendizagem. Por outro lado, Thurstone, contrário à teoria do fator *g* da inteligência, propôs sete fatores relacionados às habilidades cognitivas, considerados primários e independentes entre si, que foram denominados Fator V (compreensão verbal), Fator W (fluência verbal), Fator N (aptidão numérica), Fator S (aptidão espacial), Fator R (raciocínio), Fator P (velocidade perceptiva) e Fator M (memória) (Almeida, 1994, Ribeiro, 1998).

No ano de 1942, o psicólogo americano Raymond Cattell, com base na investigação das correlações entre as teorias de Spearman e Thurstone, concebeu a existência de dois fatores gerais da inteligência. Poucos anos depois, o psicólogo John Horn, utilizando-se de novos métodos de análise fatorial, confirmou os resultados das pesquisas de Cattell, cujos fatores gerais da inteligência foram denominados inteligência fluída (*Gf* – *fluid*

intelligence) e inteligência cristalizada (*Gc – crystallized intelligence*) (Almeida, 1994; Paro, 2006; Schelini, 2006; Vendramini & Noronha, 2003).

De maneira geral, a *Gf* refere-se aos componentes não verbais, de pouca dependência de aspectos considerados culturais e mais determinada por aspectos biológicos. Dentre suas características, pode-se dizer que, a *Gf* pressupõe a capacidade de raciocínio (processamento cognitivo) que implica capacidade de relacionar ideias, na formação de conceitos abstratos e na derivação de implicações lógicas a partir de regras gerais impostas a situações, cujo conhecimento não foi previamente memorizado (McGrew, 2004; Schelini, 2006).

Já a *Gc* está associada às capacidades exigidas na solução de grande parte dos problemas cotidianos. Seu desenvolvimento se dá por meio de experiências culturais e educacionais, sendo que, está presente na maioria das atividades escolares, muito embora não possa ser entendida como sinônimo de desempenho escolar (Carroll, 1993; McGrew, 2004; Schelini, 2006). Segundo os autores Ackerman e Rolfhus (1999), Carroll (1993) e Schelini (2006), a inteligência cristalizada evolui de acordo com o aumento etário enquanto que, a inteligência fluída, parece sofrer um declínio com o avanço etário devido, principalmente, à degeneração gradual das estruturas fisiológicas.

Em 1965, considerando os dois fatores da inteligência somados a novas pesquisas, Horn aperfeiçoou e expandiu a teoria indicada por Cattell, em que pôde organizar a inteligência em nove habilidades intelectuais. Considerando a organização proposta por Horn, as habilidades intelectuais se dividem em Inteligência Fluída, Inteligência Cristalizada, Processamento Visual, Processamento Auditivo, Memória de Curto Prazo, Memória de Longo Prazo, Rapidez de Processamento, Rapidez de Decisão e Habilidade

Quantitativa. Posteriormente, Horn, acrescentou a habilidade de compreensão em leitura e escrita, totalizando dez fatores da inteligência ou, dez habilidades intelectuais (McGrew, 2004; Schelini, 2006; Nogueira, 2008).

Carroll (1993) defende que o modelo Cattell-Horn é o que parece oferecer a mais bem fundamentada e razoável aproximação a uma teoria aceitável da estrutura cognitiva das habilidades. Assim, diante dos resultados das pesquisas realizadas por Cattell e Horn, Carroll, utilizando-se do modelo acima citado, pôde aperfeiçoá-lo culminando na Teoria dos Três Estratos da Inteligência, conforme já referenciado. Na Tabela 1 é descrito os dez fatores amplos do Estrato II e os seus fatores específicos relacionados.

Tabela 1

Descrição dos dez fatores amplos do modelo CHC e habilidades relacionadas

Estrato II	Descrição	Habilidades Específicas – Estrato I
<i>Gf</i> – Inteligência Fluída	Operações mentais de raciocínio associadas à capacidade de resolver problemas que independem de conhecimentos previamente adquiridos.	Relacionamento de idéias, indução de conceitos abstratos, compreensão de implicações, extrapolação e reorganização de informações.
<i>Gc</i> – Inteligência Cristalizada	capacidade de raciocínio adquirida por meio de repertório de conhecimentos adquiridos em experiências de aprendizagem.	Desenvolvimento da linguagem, conhecimento lexical, habilidade de discriminar sons, informação verbal geral, informação sobre cultura, habilidade de comunicação, produção oral e fluência, sensibilidade gramatical, proficiência e aptidão em língua estrangeira.
<i>Gq</i> – Conhecimento Quantitativo	Refere-se ao estoque de conhecimentos declarativos e de procedimentos quantitativos.	Utilização de informação quantitativa e manipulação de símbolos numéricos. Conhecimento matemático.
<i>Grw</i> – Leitura e Escrita	Conhecimento adquirido a partir da compreensão de textos e expressão da escrita.	Decodificação da leitura, compreensão da leitura e linguagem verbal, habilidade de cloze, habilidade de soletração, habilidade de escrita, conhecimento e utilização de idiomas.
<i>Gsm</i> – Memória de Curto Prazo	capacidade associada à manutenção de informações na consciência por um curto espaço de tempo e com poder de recuperá-las logo em seguida.	Memória de dígitos e memória de trabalho.
<i>Gv</i> – Processamento Visual	Associado aos diferentes aspectos do processamento de imagens (geração, transformação, armazenamento e recuperação).	Visualização, relações espaciais, velocidade e flexibilidade de conclusão, memória visual, exploração do espaço, percepção integrada de séries, estimação de comprimento, ilusão perceptual, percepção de alterações e manipulação de imagens.
<i>Ga</i> – Processamento Auditivo	Refere-se à capacidade de gerar, perceber, armazenar, analisar, manipular e transformar sonoridades.	Codificação fonética, discriminação geral dos sons da fala, memória para padrões sonoros, codificação temporal, discriminação musical, intensidade sonora, frequência de discriminação sonora, escuta e discurso de fatores liminares, padrões absolutos e localização de sons.

Tabela 1

Descrição dos dez fatores amplos do modelo CHC e habilidades relacionadas

(continuação)

Estrato II	Descrição	Habilidades Específicas – Estrato I
<i>Glr</i> – Armazenamento e Recuperação da Memória de Longo Prazo	capacidade associada à extensão e fluência em que as informações ou conceitos são recuperados pela memória de longo prazo por meio de associação.	Memória associativa, memória significativa, recordação livre da memória, fluência associativa, fluência expressiva, facilidade de nomeação, fluência verbal, fluência figural, flexibilidade figural, sensibilidade de resolução de problemas, originalidade/criatividade e habilidade de aprendizagem.
<i>Gs</i> – Velocidade de Processamento	Refere-se à capacidade automática e fluente na resolução de tarefas cognitivas, especialmente quando é necessário o uso da atenção e da concentração (automatismos mentais).	Rapidez perceptual, rapidez de performance, habilidade numérica, rapidez de raciocínio, fluência na leitura e fluência na escrita.
<i>Gt</i> – Rapidez de Decisão	Reação à tomada de decisão rápida em resposta a estímulos mais simples.	Tempo de reação, tempo de escolha, velocidade.

Fonte: Cattell-Horn-Carroll (CHC) Definition Project, de K. S. McGrew (2004), adaptada por Nogueira (2008, p. 16)

No que se refere ao desempenho acadêmico, Flanagan, McGrew e Ortiz (2000), descrevem os quatro fatores das habilidades cognitivas que estariam mais associadas ao contexto, o fator *Gq* (Conhecimento Quantitativo), o fator *Grw* (Leitura e Escrita), o fator *Gf* (Inteligência Fluída) e o fator *Gc* (Inteligência Cristalizada). Contudo, segundo os autores, a Inteligência Cristalizada é o fator que mais contribui tanto para as habilidades cognitivas quanto para as habilidades acadêmicas, dado ao fato de que a *Gc* é um fator intermediário entre os processos gerais do pensamento, o desenvolvimento da linguagem e os processos gerais concernentes à aprendizagem.

Ackerman e Rolfhus (1999) enfatizam que alguns pesquisadores impõem o bom desempenho na resolução de tarefas e tomadas de decisões aos fatores *Gc* e *Gf* – para esses

autores, esse modo de se pensar a inteligência é uma maneira pobre de se avaliar a abrangência intelectual em adultos. Diante desta premissa, Rolfhus e Ackerman (1999) propuseram uma teoria da inteligência que introduz o domínio do conhecimento específico como parte integrante da inteligência em adultos. Esse conhecimento específico exerce importante papel no desempenho intelectual e pode ser atribuído, principalmente, aos níveis mais elevados e integrados do conhecimento (Rolfhus & Ackerman, 1999).

Seguindo essa corrente teórica, McGrew (2008) aponta uma lacuna significativa no modelo CHC e apresenta a inclusão de, pelo menos, seis grandes fatores ou domínios referentes as habilidades cognitivas que melhor explicariam o modelo. Sendo assim, o Estrato II seria acrescido das seguintes habilidades, *Gkn* (conhecimento específico), *Gh* (habilidade tátil), *Gk* (habilidade sinestésica), *Go* (habilidade olfativa), *Gp* (habilidade psicomotora) e, por último, *Gps* (velocidade psicomotora).

A descrição dos fatores *Gh*, *Gk*, *Go*, *Gp* e *Gps* não será privilegiada neste trabalho mas, um domínio particular se faz importante para a presente pesquisa, trata-se do fator *Gkn*. Como já citado, *Gkn* refere-se ao domínio do conhecimento específico e reflete, fortemente, domínios especializados do conhecimento desenvolvido por um período prolongado de tempo e que exige intensa e sistemática prática e formação, cuja manutenção é motivada pela regularidade e interesse, de acordo com as diferenças individuais de cada sujeito (Ackerman & Lohman, 2006; McGrew & Evans, 2004; McGrew, 2008).

Rolfhus e Ackerman (1999) indicam que o conhecimento específico está diretamente associado ao *Gc*. A principal distinção entre *Gc* e *Gkn*, conforme Ackerman e Lohman (2006) e McGrew e Evans (2004), reside no fato de que a *Gc* reflete, principalmente, o conhecimento geral acumulado por meio da experiência do sujeito em

determinada cultura enquanto que, a *Gkn* representa a amplitude e a profundidade do conhecimento em domínios específicos. Assim, o fator *Gkn* estaria mais associado ao bom desempenho em, por exemplo, testes e provas acadêmicas.

A partir da sucinta descrição das teorias da inteligência, desde o conceito do fator *g* de Spearman, seguido dos sete fatores da inteligência proposto por Thurstone, dos estudos de Cattell que resultaram na teoria *Gc* e *Gf*, dos avanços de Horn ao sugerir a existência de, aproximadamente, dez fatores responsáveis pelas habilidades cognitivas, o aperfeiçoamento dessas ditas teorias que culminaram no modelo CHC de Carroll e, por fim, o acréscimo de mais seis habilidades cognitivas apresentadas por McGrew (2008) que tais estudiosos estabeleceram parâmetros para as medidas do construto em questão. Desde então, o conceito de habilidade pôde ser melhor averiguado e estendido para o contexto educacional, principalmente, quando se trata das habilidades e competências necessárias para a realização de testes de desempenho acadêmico.

Uma ressalva é feita por Primi e cols. (2001) no que concerne ao termo habilidade utilizado por Carroll. Para os autores, Carroll, ao definir o termo habilidade, o fez de forma muito ampla e generalista, cuja principal função seria determinar as variações individuais do potencial dos sujeitos na realização de um conjunto definido de tarefas. Essa ideia de habilidade pressupõe facilidade em relacionar informações indicando um potencial de realização para determinadas tarefas e conteúdos. Para os autores (2001, p. 155), “o investimento deste potencial em experiências de aprendizagem pode conduzir à maestria, que envolve um conhecimento organizado sobre um determinado tema. A ideia de maestria implica que a realização atingiu um determinado nível a partir do qual se diz que uma competência foi adquirida”. Assim, Vendramini (2000) atenta para o fato da importância

do uso de palavras específicas e conceitos claramente definidos ao utilizar-se de concepções alternativas dos conceitos de habilidade e competência.

Um conceito de habilidade que se assemelha aos propósitos deste estudo é o apresentado por Brito e cols. (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 e cols. (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 idéia 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 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 (Primi & cols., 2001). Desta maneira, alguns fatores estão associados à aquisição de competências, ao desenvolvimento de habilidades, ao montante de investimento determinado pelo sujeito e à qualidade das experiências de aprendizagem (Primi & cols., 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 Exame Nacional de Desempenho dos Estudantes (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 e cols. (2001) destacam que as instituições educacionais 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, atualmente, 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 envolvimento da família no processo de aprendizagem, situação sócio-econômica, 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.

Ponderando o desempenho acadêmico, a avaliação educacional e a qualidade dos instrumentos de medida utilizados, focos do presente trabalho, o Capítulo a seguir trata, principalmente, da instauração de programas de avaliação educacional. É também nesse *locus* que serão apresentadas, as contribuições desses programas educacionais às medidas de desempenho acadêmico, bem como as metodologias utilizadas na questão da avaliação em larga escala.

CAPÍTULO 2

AVALIAÇÃO DO ENSINO SUPERIOR

Com a emergente necessidade de um mercado educacional globalizado que privilegiasse a ampliação do acesso à educação e a diminuição das assimetrias sociais, a partir da década de 1980, os países industrializados e latino-americanos se propuseram à reflexão acerca das reformas em seus sistemas de Educação. O intuito dessa nova forma de pensar a questão educacional angaria recursos técnicos e informativos que viabilizam a distribuição adequada de recursos públicos, bem como possibilitam a avaliação da qualidade da educação ofertada à população (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – Inep, 2004).

Alguns modelos americanos e europeus serviram de inspiração para a criação e desenvolvimento dos sistemas de avaliação da educação superior vigentes na atualidade. Cabe ressaltar que os modelos anglo-americanos aludem à uma avaliação predominantemente quantitativa que visa, como objetivo último, a produção de resultados classificatórios. Já os modelos holandeses e franceses promulgam a combinação de dimensões qualitativas e quantitativas, cuja ênfase recai sobre a avaliação institucional e a análise das variáveis que compreendem essas dimensões. Nas últimas três décadas, os modelos de avaliação do Ensino Superior no Brasil, seguem, assim como os sistemas de avaliação dos países desenvolvidos, uma ou outra dessas orientações (Inep, 2004).

O mais antigo modelo de avaliação da educação superior brasileira foi proposto pela CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e é desenvolvido desde 1976 tendo como objetivo principal avaliar a qualidade dos cursos e programas de

pós-graduação. Essas ações serviram, também, de base para a promoção e implementação de novas propostas que seguem a orientação do desenvolvimento político, econômico e social do país. Contudo, até o início da década de 1980, essas discussões eram pouco motivadas e, assim, alvo de pouco destaque (Inep, 2004).

Desde então, o panorama da Educação Superior brasileira vem se consolidando como uma área que requer atenção quando diz respeito à qualidade de ensino ofertada no país, bem como no que se refere ao processo ensino-aprendizagem concernentes aos estudantes nesse contexto. A preocupação com a avaliação do ensino superior torna-se mais intensa a partir do ano de 1983, com a instituição, por meio do Ministério da Educação (MEC), do Programa de Avaliação da Reforma Universitária (PARU), cujos propósitos eram aferir a disseminação e produção do conhecimento, além de avaliar a gestão das Instituições de Ensino (Polidori, Marinho-Araújo & Barreyro, 2006).

Para tanto, criaram-se novos programas governamentais que substituíam os já em vigor e com finalidades semelhantes, “Comissão de Notáveis” em 1985 e, em 1986, o Grupo Executivo da Reforma da Educação Superior (GERES). Já na transição da década de 1980 para 1990, algumas Instituições de Ensino (IES), preocupadas com a qualidade do ensino, instituíram modelos próprios de avaliação interna para seus cursos de graduação. Tais avaliações despertaram o interesse, já assíduo, do MEC e serviram de subsídio para a elaboração, em 1993, do Programa de Avaliação Institucional das Universidades Brasileiras – PAIUB, como uma primeira tentativa de implantação de um sistema nacional de avaliação do ensino superior (Polidori & cols., 2006).

O PAIUB surge com as inquirições respaldadas na participação voluntária das universidades, ideando um processo principiado com a auto-avaliação que englobava toda a

instituição e concluíam-se com a avaliação externa. Constituiu novos métodos de relacionamento para lidar com o conhecimento e a formação, consolidando caminhos almejados, tecendo as perspectivas em concordância com a comunidade acadêmica e a sociedade (Inep, 2004).

Apesar de sua breve atuação, deu autenticidade à tradição da avaliação e impulsionou transformações relevantes na dinâmica acadêmica. Seu encadeamento foi afetado pelo fim do apoio do MEC, mesmo tendo adquirido ampla participação das universidades brasileiras, tornando-se um processo de avaliação meramente interno às instituições, o que ocasionou repercussões negativas, afetando o seu desenvolvimento (Inep, 2004).

Diante de todas essas tentativas de criação de um mecanismo que contribuísse efetivamente para a questão da avaliação das IES, em novembro de 1995, surge o Conselho Nacional de Educação (CNE) que possibilitou maior abertura para que um conjunto de avaliações periódicas fossem submetidas às instituições e cursos superiores (Brasil, 1995). Assim, baseados em conteúdos mínimos de cada curso, exames anuais foram elaborados com o escopo único de avaliar os estudantes em fase de conclusão de curso acerca das competências adquiridas e os conhecimentos oriundos da graduação (Paiva, 2008).

O Exame Nacional de Cursos (ENC) surge no ano de 1996 com o principal objetivo de avaliar o desempenho acadêmico de estudantes universitários. Conhecido por “Provão”, o ENC era um exame de âmbito nacional e fazia parte de um sistema maior de avaliação composta por outros modelos avaliativos como, por exemplo, a Avaliação das Condições de Ensino (ACE) e a Avaliação para credenciamento de IES privadas, com o intuito de

equalizar a qualidade de ensino nas universidades brasileiras, fossem elas públicas ou privadas (Polidori & cols., 2006).

O ENC-Provão sofreu diversas críticas devido, principalmente, ao seu caráter avaliativo que contemplava apenas os estudantes concluintes dos cursos de graduação (Paiva, 2008). Nessa perspectiva, Soares, Ribeiro e Castro (2001) enfatizam que os problemas metodológicos tais como o fato de não levar em consideração a situação sócioeconômica dos estudantes, bem como seu conhecimento prévio, ou seja, sua situação na condição de ingressantes no Ensino Superior, foram agentes motivadores de discussões acerca da ineficácia do exame.

Com o objetivo de promover melhorias na avaliação do Ensino Superior, em abril de 2004, foi instituído um novo modelo, o Sistema Nacional de Avaliação da Educação Superior – SINAES (Paiva, 2008). De acordo com Limana e Brito, na ocasião de sua implementação (2005, p. 11), “...todos os elementos da avaliação, no contexto do SINAES, são essencialmente educativos. Trata-se de um projeto, em movimento, dinâmico; examina o passado e julga o presente”.

O objetivo do SINAES é averiguar, em âmbito nacional, a qualidade das instituições de nível superior, dos cursos de graduação das diversas áreas do conhecimento, o desempenho dos estudantes ingressantes e concluintes e, ainda, o perfil socioeconômico desses estudantes. O Exame Nacional de Desempenho dos Estudantes (ENADE) é um dos instrumentos que integram o SINAES e, por ser um exame de abrangência nacional e realizado por processos estatísticos de amostragem aleatória, o ENADE torna possível inferir sobre o cenário atual dos cursos avaliados por esse sistema (MEC, 2004).

Sendo assim, segundo o MEC (2004, p. 22),

O ENADE verificará o desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares dos respectivos cursos de graduação, suas habilidades para ajustamento às exigências decorrentes da evolução do conhecimento e suas competências para compreender temas exteriores ao âmbito específico de sua profissão ligados à realidade brasileira e mundial e a outras áreas do conhecimento.

O ENADE foi realizado pela primeira vez em Novembro de 2004 e considerou 14 áreas de conhecimento: Agronomia, Educação Física, Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Medicina, Medicina Veterinária, Nutrição, Odontologia, Serviço Social, Terapia Ocupacional e Zootecnia. Os estudantes que participaram do exame foram selecionados por meio de procedimentos amostrais entre estudantes ingressantes e concluintes dos cursos de graduação. Em 2005, com um aumento de cerca de 43% de cursos abrangidos, foram avaliadas as áreas de: Arquitetura e Urbanismo, Biologia, Ciências Sociais, Computação, Engenharias (modalidades distintas), Filosofia, Física, Geografia, História, Letras, Matemática, Pedagogia e Química, totalizando 20 grupos distintos (Inep, 2004).

A investigação proposta pelo ENADE conta com a utilização de quatro instrumentos, a saber, a prova, composta por itens que pretendem aferir o desempenho de estudantes com relação aos conteúdos programáticos a que se inserem, bem como conhecimentos em outras áreas; o questionário de impressões sobre a prova, que tem por objetivo verificar a percepção do estudante referente a aspectos específicos da prova como, por exemplo, nível de dificuldade, tempo de realização, entre outros; questionário de perfil do estudante cujo propósito é compor o perfil do estudante, com informações acerca de

suas vivências dentro da IES; e, o questionário do coordenador do curso, que se propõe a buscar informações dos profissionais sobre aspectos da prova e suas relações com o projeto pedagógico, além das condições gerais de ensino em seus cursos de graduação (MEC, 2005).

Os resultados desses exames são divulgados anualmente pelo MEC e servem de índices considerados confiáveis da Educação Superior brasileira. Um dos motivos que pressupõe inferências seguras desse panorama diz respeito à composição da amostra que, até o ano de 2008, era selecionada aleatoriamente por amostragem estratificada e representativa da população em consideração. A partir de 2009, todos os estudantes que preenchem o critério estipulado para ingressantes e concluintes, indicados por suas IES, são convocados a participarem do exame. Na Tabela 2 são apresentados os resultados divulgados pelo MEC em que é possível visualizar a distribuição dos estudantes por cursos participantes do ENADE 2005.

Tabela 2

Distribuição dos estudantes e cursos submetidos ao ENADE em 2005

Cursos	Ingressantes	Concluintes	Total	%
Arquitetura e Urbanismo	3.621	2.912	6.533	02,4
Biologia	13.542	10.933	24.475	08,8
Ciências Sociais	1.680	2.150	3.830	01,4
Computação	18.004	14.056	32.060	11,6
Engenharia – Grupo I	4.562	4.163	8.725	03,1
Engenharia – Grupo II	10.154	7.012	17.166	06,2
Engenharia – Grupo III	2.598	1.878	4.476	01,6
Engenharia – Grupo IV	2.954	2.185	5.139	01,9
Engenharia – Grupo V	712	449	1.161	00,4
Engenharia – Grupo VI	4.751	2.164	6.915	02,5
Engenharia – Grupo VII	2.639	779	3.418	01,2
Engenharia – Grupo VIII	1.412	923	2.335	00,8
Filosofia	2.363	2.056	4.419	01,6
Física	2.574	1.654	4.228	01,5
Geografia	5.668	6.075	11.743	04,2
História	9.537	9.075	18.162	06,5
Letras	23.330	22.870	46.200	16,7
Matemática	10.229	9.243	19.472	07,0
Pedagogia	23.318	26.179	49.497	17,8
Química	3.952	3.120	7.072	02,5
Total	147.600	129.876	277.476	100,0

Fonte: Relatório Técnico – ENADE/2005, disponível em www.inep.gov.br/enade.

Com a constituição do ENADE como um componente do modelo atual de avaliação da Educação Superior e de acordo com as exigências do SINAES, a demanda por avaliações institucionais nas universidades privadas teve um acréscimo e tornou-se um

fértil campo para pesquisas no âmbito educacional. Os índices gerados pelo exame apresentam-se como valores passíveis de comparação, fornecendo dados e estimativas nacionais que contribuem para o diagnóstico da situação vigente nas IES.

Contudo, ao longo dos oito anos de exercício, o ENADE angariou críticas quanto à sua proposta inicial. Nesse sentido, Verhine, Dantas e Soares (2006) enumeram alguns entraves referentes ao ENADE como instrumento de avaliação de cursos dentre eles, a falta de especificidade no que se refere as comissões técnicas responsáveis pelo processo e formatação da prova; a insuficiência do número de questões da prova que não permite maior abrangência avaliativa das várias etapas do curso; e, ainda, a ausência da pré-testagem dos itens que, caso ocorresse, dispor-se-ia de critérios que viabilizassem maior qualidade à esses itens.

Seguindo a discussão sobre as críticas ao ENADE, Brito (2008) chama atenção para o fato de que, a partir de 2008, o MEC passou a divulgar um *ranking* das notas da prova. Para a autora, os objetivos concebidos para a aplicação do Exame se perderam, já que ao se promover esse *ranqueamento*, os ingressantes passam a “atrapalhar” os resultados das IES e, assim, perde-se o foco da aferição do progresso dos estudantes que deve ter por elemento norteador o potencial de aprendizagem (desempenho dos ingressantes) e o domínio da área e competências profissionais (desempenho dos concluintes).

Ainda sobre a questão de *ranqueamento* de cursos das IES, Dias Sobrinho (2010) menciona que a divulgação de *rankings* por meio do MEC e da mídia, torna o ENADE um instrumento isolado, desconsiderando o Sistema de Avaliação como um todo. Dessa maneira, o autor menciona a supervalorização dessas informações em que,

(...) a mídia acabou elaborando equivocadamente *rankings*, tomando como base apenas o ENADE e em uma de suas aplicações, isoladamente, o que já é uma deturpação desse instrumento. Muitas IES se acostumaram a promover ampla propaganda dos itens positivamente avaliados, ainda que parciais e destacados do contexto, para melhorarem sua imagem pública, e, com isso, obterem vantagens competitivas no mercado educacional. Contudo, o ENADE não pode ser destacado do conjunto dos instrumentos do SINAES nem ser considerada uma única nota, separadamente. Exceto no que se refere à formação geral e no mesmo ano, os resultados de diferentes cursos não são homologáveis, pois as provas são distintas e as matrizes elaboradas por diferentes comissões de especialistas (Dias Sobrinho, 2010, p. 215).

Primi e cols. (2009) e Primi e cols. (2010) relatam algumas defasagens do ENADE no que concerne ao valor agregado, que pode ser compreendido como uma medida dos ganhos acadêmicos dos alunos em sua trajetória educacional. Primi e cols. (2009) ampliam as possibilidades de reflexão sobre a questão do quanto a qualidade de um curso pode ser avaliada por meio do desempenho dos estudantes. Para os autores (2009), a limitação principal do ENADE refere-se ao tipo de delineamento empregado no Exame que não é capaz de oferecer, satisfatoriamente, a medida das mudanças ocorridas nos estudantes ao longo da trajetória acadêmica. Para suprir essa deficiência seriam necessárias mensurações longitudinais e a equivalência dos grupos ao início do curso para que, dessa maneira, as interpretações dos resultados do ENADE possam ser atribuídas como um indicador eficiente da qualidade dos cursos. Contudo, o ENADE utiliza-se de dados transversais para obtenção desses valores, isto é, as medidas dos alunos ingressantes e concluintes no mesmo

ano. Desse modo, os resultados dessas medidas contemplam o curso em detrimento do aluno e, além disso, essa característica do Exame pode limitar as interpretações desses resultados e suscitar questionamentos quanto à metodologia da prova.

As avaliações internas nas IES servem de ferramenta não somente para esclarecer a condição em que se encontram os cursos, os estudantes, os docentes e as demais variáveis inerentes ao contexto acadêmico, mas também como um modo de angariar recursos científicos que possam servir de componente de informação, formação e execução de procedimentos que visem a melhoria da educação em sua totalidade. Para tanto, os instrumentos devem seguir o crivo e as especificidades respectivas à qualidade da avaliação pretendida, com itens que demonstrem, realmente, um propósito a ser avaliado.

O caráter avaliativo no ensino superior, conforme Muñiz e Fonseca-Pedrero (2008), tem por finalidade a elaboração de um diagnóstico rigoroso que permita melhorar os três grandes parâmetros que determinam a qualidade de um IES, são eles, a investigação, a docência e a gestão. Sem uma avaliação precisa desses três parâmetros não se pode realizar um diagnóstico rigoroso que permita gerar planos de melhorias embasados em dados empíricos.

Goldstein (1998) destaca a importância da avaliação e das pesquisas em Educação já que a pesquisa educacional não trata, simplesmente, dos papéis desempenhados pelas instituições e professores mas, também, sobre a aprendizagem, sobre o papel da sociedade e da cultura em aprendizagem e, ainda, sobre a inter-relação entre todos esses fatores. Para o autor é evidente que não pode haver nenhuma garantia de que a pesquisa em educação, como em qualquer outra área, será de alta qualidade. E, nesse caso, o melhor a ser feito é

colocar em contínua prática procedimentos que investiguem os atributos dos itens e testes utilizados em Educação.

Para Raudenbush (2005), a eficácia de pesquisas na área educacional, considerando a qualidade dos testes e da metodologia para análise dos resultados, promove propostas de intervenção que podem incluir novos currículos, novas tecnologias, melhores formas de preparação de docentes e capacitação dos serviços prestados pelas IES. Dessa forma, os resultados dessas avaliações apresentam o contexto atual do setor, bem como, norteiam medidas a serem tomadas.

Bird e cols. (2005) defendem alguns critérios que devem ser monitorados quando se pretende uma avaliação abrangente e com rigores científicos na análise dos dados obtidos. Para os autores, em uma avaliação educacional, a coleta dos dados deve ser especificada para cada indicador de desempenho, isso incluirá a escolha das fontes de dados, que pode se basear em uma amostra ou um censo. A qualidade dos dados também deve ser privilegiada, já que uma pequena ou grande quantidade de dados defeituosos pode causar um viés à pesquisa, haja vista que não há nenhuma segurança em termos de quantidade em detrimento da qualidade. Seguindo a lógica coleta-análise-interpretação dos dados, os autores argumentam que o esforço despendido na análise e interpretação dos dados deve ser proporcional ao esforço consumido na coleta dos mesmos.

De acordo com alguns pesquisadores (Goldstein, 1998; Raudenbush, 2005; Bird & cols., 2005; Muñiz & Fonseca-Pedrero, 2008), a avaliação educacional conta com certas deficiências e limitações. Para Muñiz e Fonseca-Pedrero (2008), a principal limitação refere-se à carência de um modelo geral que integre e guie as avaliações. Para os autores, um modelo deve dar respostas claras, ao menos, à sete questões chaves, a saber, o que se

está avaliando, quais são as partes legitimamente implicadas na avaliação, quem se avalia, como se avalia (que método será utilizada), que *feedback* pode ser oferecido, planos de melhora gerados por meio da avaliação e a opinião das partes implicadas na avaliação.

Considerando os fatores que regem uma avaliação eficaz, deve-se sopesar e definir de forma clara, concisa e operativa aquilo que se deseja avaliar. Torna-se necessário um compromisso entre a importância do construto a ser medido e a objetividade dos instrumentos de medida utilizados. No contexto da avaliação universitária, conforme Muñiz e Fonseca-Pedrero (2008), existem numerosos objetivos de avaliação; quase todos eles podem ser classificados em três grandes blocos, indivíduos, resultados e sistemas. Avaliam-se estudantes, professores, administração, serviços oferecidos pelas IES e seus gestores. No que se refere aos resultados, a gama é bastante ampla, destacando-se os projetos de investigação, os artigos, as teses e os livros publicados, bem como o currículo e o plano de estudo da instituição em avaliação. Já a avaliação dos sistemas alude às faculdades, aos institutos, aos grupos de investigação, bibliotecas e universidades.

Os propósitos dessas avaliações necessitam ser condizentes à legitimidade das medidas que, por seu lado, devem ser bem estabelecidas por um conjunto de estudos planejados a fim de testar se o conteúdo do instrumento atinge os fins delineados. Tais estudos testam a plausibilidade, não somente, do instrumento mas, também, a qualidade das informações e interpretações pretendidas com a utilização do instrumento.

Histórico da autoavaliação de uma universidade particular do interior de São Paulo

A qualidade do ensino ofertado à população, desde sempre, é uma das preocupações das IES. As primeiras experiências de autoavaliação da universidade em questão mencionam a aderência ao PAIUIB, ainda na década de 1990, servindo de base para novas maneiras de se pensar a qualidade dos cursos de graduação, bem como, promover um diagnóstico acerca dos serviços prestados pela universidade. Já no início do ano 2000, foi criada, de acordo com o Estatuto da universidade, a Pró-Reitoria de Desenvolvimento Institucional com o objetivo de definir e traçar diretrizes acerca das políticas de avaliação institucional. No ano de 2004 foi constituída a Comissão Própria de Avaliação (CPA), uma das exigências do SINAES, que se propõe, entre outros encargos, à responsabilidade pela direção dos processos de avaliação interna (<http://www.usf.edu.br/cpa/>).

Em termos gerais, os objetivos da CPA consistem em buscar compreender o significado e o alcance das atividades propostas, sistematizar informações, investigar e diagnosticar as fragilidades e potencialidades dos serviços prestados e promover estratégias de melhorias. Para tanto, a CPA sugere as dimensões para uma avaliação eficaz, quais sejam, a missão e o Plano de Desenvolvimento Institucional (PDI); as políticas de ensino que contemplem o ensino, a pesquisa, a pós-graduação e a extensão; a responsabilidade social; o diálogo entre instituição e sociedade; as relações de carreira de seus funcionários; a organização e a gestão da instituição; os recursos e a infraestrutura física; o atendimento aos estudantes e ex-estudantes; e, por fim, a sustentabilidade financeira (<http://www.usf.edu.br/cpa/>).

Tomando como recorte o que concerne a avaliação discente, uma das funções da CPA, consiste em avaliar a questão da relação ensino-aprendizagem, bem como o desempenho dos estudantes em diferentes anos de graduação; pensando nisso, no ano de 2005 criou-se a Prova Pluridisciplinar. Esta prova era composta por questões elaboradas pelos próprios docentes do curso e tinha por objetivo avaliar o desempenho dos estudantes, de acordo com o conteúdo abrangido por cada disciplina. Os resultados obtidos por meio da aplicação da prova Pluridisciplinar informavam acerca do desempenho dos estudantes, mas somente permitia a comparação desse desempenho dentro da própria amostra. Para sanar essa deficiência, em 2008 foi criado o Programa de Avaliação Continuada (PAC) (<http://www.usf.edu.br/cpa/>).

O PAC é um programa de avaliação discente, cujo objetivo é aferir as habilidades acadêmicas, isto é, o desempenho dos estudantes no que concerne ao conhecimento adquirido ao longo da graduação, o desenvolvimento e aquisição das competências profissionais, a formação geral e profissional desses estudantes, bem como suas relações com a realidade do país. Seus princípios norteadores seguem os mesmos propósitos do ENADE, propósitos esses, que viabilizam um diagnóstico da situação acadêmica dos estudantes, no que se refere ao aprendizado na condição de graduandos (<http://www.usf.edu.br/cpa/>).

Para que esses propósitos fossem cumpridos a CPA contou com o auxílio do Laboratório de Avaliação Psicológica e Educacional (LabAPE) da Universidade São Francisco (USF). O LabAPE é vinculado ao Programa de Pós-Graduação *Stricto Sensu* em Psicologia da USF e contempla duas, das três, linhas de pesquisa propostas pelo Programa a saber, “Construção, Validação e Padronização de Instrumentos de Medida” e “Avaliação

em Psicologia Educacional”. Cabe salientar os objetivos deste Laboratório que implicam no desenvolvimento e no aprimoramento de instrumentos e técnicas de medição nas áreas da Psicologia e da Educação; na investigação das aplicações práticas de instrumentos nos contextos educacionais; no desenvolvimento das aplicações da Psicometria e da Teoria de Resposta ao Item (TRI); no desenvolvimento de recursos tecnológicos de informática que possam auxiliar no tratamento dos dados; e, finalmente, na organização das informações coletadas (<http://www.labape.com.br/labape/objetivos.html>).

O LabAPE também está vinculado 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 (PACES/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. 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 Ministério da Educação (Primi, 2006; <http://www.labape.com.br/labape/projetos.html>).

Assim, desde 2008, cabe ao LabAPE a tarefa de analisar as provas do ENADE; selecionar itens com boas propriedades psicométricas por meio da TRI; construir e analisar as provas do PAC; comparar, por meio da criação de uma escala de medidas comum entre o

ENADE e o PAC (equalização), o desempenho dos estudantes da universidade em questão ao desempenho dos estudantes brasileiros; e, elaborar relatórios com os resultados provenientes desta tarefa – conforme descrito na presente pesquisa.

Nessa perspectiva, os testes educacionais em larga escala como o ENADE, pretendem aferir o desempenho, isto é, o percurso acadêmico dos estudantes adquirido na situação de graduandos (Primi, 2002; Primi, Santos, & Vendramini, 2002). Para que sejam realizadas inferências válidas e confiáveis do desempenho acadêmico com relação aos resultados obtidos pelo ENADE e pelo PAC, as provas deverão apresentar itens consonantes com a realidade que pretendem avaliar, ou seja, ser bem construída e averiguar, de fato, o conteúdo e os conceitos propostos para cada área de atuação. Assim, métodos e modelos, como os descritos no Capítulo a seguir, que possibilitem análises confiáveis devem ser utilizados, considerando a importância da avaliação educacional no contexto social, educacional e econômico do país.

CAPÍTULO 3

A TEORIA DE RESPOSTA AO ITEM

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 há pouco, já que teve seu desenvolvimento elaborado por diversos autores a partir dos anos 50, 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, para melhor explicar os resultados obtidos é necessário estimar os valores destes parâmetros, com base nas respostas dos sujeitos aos itens, de acordo como os modelos de três, dois e um parâmetros. A esse processo é dado o nome de calibração ou parametrização, e é feito com o auxílio de programas específicos.

O modelo de três parâmetros da TRI inclui em sua função a probabilidade de acerto ao acaso. Esse índice indica a probabilidade de resposta correta ao item por respondentes com baixa habilidade nas dimensões. A fórmula matemática correspondente à probabilidade de um indivíduo j acertar a um item i , é dada por:

$$P(X_{ij} = 1 | \theta_j) = P_i(\theta) = c_i + (1 - c_i) \frac{1}{1 + e^{-D a_j (\theta - b_i)}} \quad (1)$$

Nesta fórmula a variável X_{ij} é uma variável dicotômica que assume os valores 1, quando o indivíduo j responde corretamente ao item i , ou 0, quando o indivíduo não responde corretamente ao item i . A probabilidade do sujeito j responder corretamente ao item i é denotada por $P_i(\theta_j)$, em que θ_j é o valor da variável latente (aptidão, proficiência ou habilidade) do sujeito j . O modelo apresentado em (1) considera três parâmetros, o parâmetro a_i que representa o índice de discriminação do item, com valor proporcional à inclinação da Curva Característica do Item (CCI) no parâmetro b_i . O parâmetro b_i corresponde ao índice de dificuldade do item, medido na mesma escala de habilidade; seu valor é correspondente à habilidade necessária para uma probabilidade de acerto igual a $(1+c_i)/2$.

O parâmetro c se refere à probabilidade de acerto ao acaso ou de sujeitos com baixa habilidade responderem corretamente ao item i . A constante e é igual a 2,718 e a constante D igual a 1, e quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva b normal, utiliza-se o valor $D=1,7$. A seguir, são apresentadas, na Figura 2, três curvas características do item que seguem o modelo de três parâmetros.

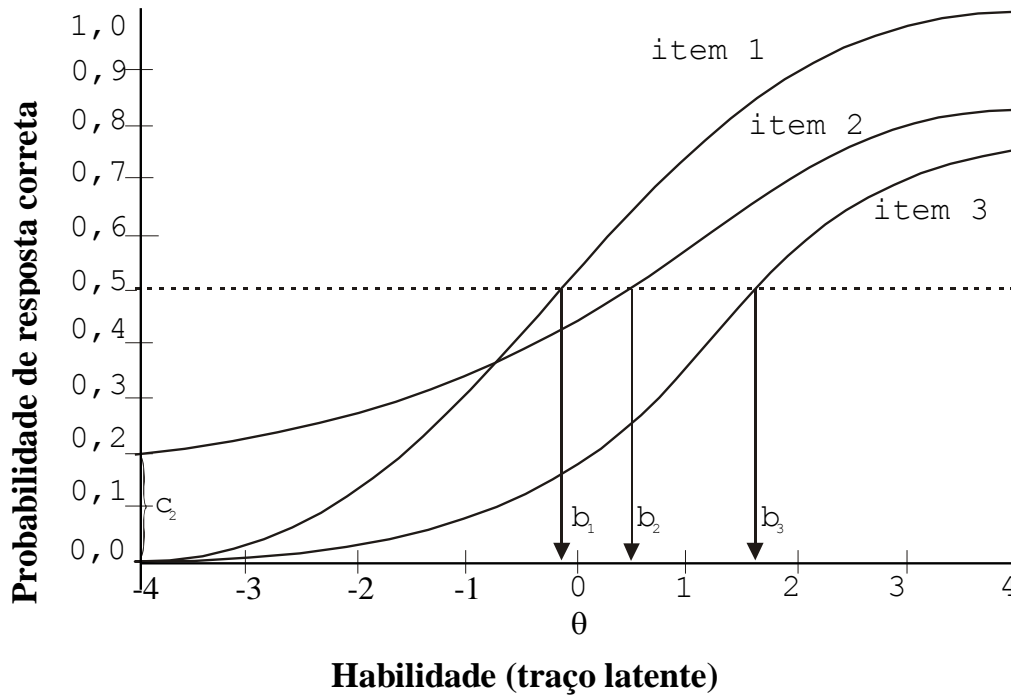


Figura 2. Modelo logístico de três parâmetros

O modelo logístico de dois parâmetros considera a probabilidade de acerto ao acaso igual a zero para todos os itens ($c=0$). Por fim, o modelo de um parâmetro considera, também, que todos os itens têm a mesma discriminação. O modelo de um parâmetro, ou modelo de Rasch, relaciona, portanto, a probabilidade de um indivíduo acertar o item com o índice de dificuldade do item, como apresentado a seguir (2):

$$P(X_{ij} = 1 | \theta_j) = P_i(\theta) = \frac{1}{1 + e^{-D(\theta - b_i)}} \quad (2)$$

As principais diferenças entre os modelos estão na forma matemática da curva característica do item e no número de parâmetros do modelo. No primeiro modelo, o único parâmetro é o índice de dificuldade do item; no segundo modelo, são dois os parâmetros – o índice de dificuldade e o índice de discriminação; e no terceiro, são três – o índice de dificuldade, o índice de discriminação e a probabilidade de acerto ao acaso. Esses modelos são expressões matemáticas que fornecem a probabilidade de resposta correta a um item como função do traço medido do indivíduo e dos parâmetros dos itens cuja discriminação corresponde à inclinação da Curva Característica do Item (CCI) e indica o quanto, indivíduos de diferentes habilidades, diferem quanto à probabilidade de acertar um item. Nos três modelos matemáticos apresentados, a probabilidade de acertar um item é dada em função da variável latente, e assume valores entre 0 e 1 (Vendramini, 2005).

Pode-se dizer que os precursores dessa dita teoria, propunham um novo modelo de análise de itens que abrangesse construtos e procedimentos que, antes, não eram esclarecidos por meio da Teoria Clássica dos Testes (TCT). A TCT, por se tratar de um modelo que é intrinsecamente dependente do objeto medido, limitava o campo de trabalho e que, de certa forma, implicava em um infértil panorama às novas descobertas.

Thurstone (1928, citado por Pasquali & Primi, 2007, p. 11) referindo-se aos problemas que limitavam a TCT assinala que

um instrumento de medida, na sua função de medir, não pode ser seriamente afetado pelo objeto de medida. Na extensão em que sua função de medir for assim afetada, a validade do instrumento é prejudicada ou limitada. Se um metro mede diferentemente pelo fato de estar medindo um tapete, uma pintura ou um pedaço de papel, então nesta mesma extensão a confiança neste metro como instrumento de

medida é prejudicada. Dentro dos limites de objetos para os quais o instrumento de medida foi produzido, sua função deve ser independente com respeito à medida do objeto.

Pasquali e Primi (2007) relatam alguns problemas que acompanham a TCT como, por exemplo, a premissa de que a dificuldade do teste varia de acordo com a amostra utilizada, ou seja, um item pode se tornar mais fácil ou mais difícil para uma dada população. Os autores ainda revelam outros entraves encontrados na TCT principalmente nos quesitos referentes à discriminação do item, fidedignidade de testes e, ainda, aos erros de medida.

Assim, para resolver os problemas que não encontram uma boa solução dentro da TCT, a TRI se apresenta como um modelo com pressupostos mais fortes e desenvolvidos no sentido de suprir a teoria clássica. Para Muñiz (2010), atualmente, a TRI domina o panorama psicométrico, dada sua funcionalidade que permite uma avaliação mais precisa e rigorosa que, por sua vez, resulta em uma intervenção eficaz. O autor destaca, na Tabela 3, as principais diferenças entre a TCT e a TRI.

Considerando as informações entre as duas teorias apresentadas na Tabela 3, pode-se dizer que um dos diferenciais mais importantes da TRI refere-se ao processo de identificação e da relação entre as propriedades dos itens e das pessoas, tendo como ponto de partida modelos probabilísticos. De acordo com Nunes e Primi (2009, p. 42) “essa característica é que proporciona ao modelo o termo ‘teoria’, uma vez que, diante de dados empíricos, que são as respostas dadas por um grupo de pessoas a um conjunto de itens, é construído um modelo que tenta explicar a relação entre eles”.

Tabela 3

Diferenças entre a Teoria Clássica e a Teoria de Resposta ao Item

Aspectos	Teoria Clássica	Teoria de Resposta ao Item
Pressupostos	Fracos	Fortes
Invariância da medição	Não	Sim
Invariância das propriedades do teste	Não	Sim
Ênfase	No teste	No item
Relação item-teste	Indeterminada	Curva característica do Item
Descrição dos itens	Índices de dificuldade e de discriminação	Parâmetros a, b, c
Erros de medida	Erros típicos de medida comum para toda amostra	Em função da informação (varia segundo o nível de aptidão)
Tamanho amostral	Pode funcionar bem com amostras entre 200 e 500 sujeitos, aproximadamente	Recomenda-se mais de 500 sujeitos, embora dependa do modelo

Fonte: adaptado de J. Muñiz (2010).

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. 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 em larga escala, a metodologia dessa teoria possibilita o melhor conhecimento do desempenho dos estudantes e das variáveis suscetíveis a esse contexto.

Em caráter de exemplo, análises utilizando-se os modelos propostos pela TRI também vêm sendo realizadas com os dados do Sistema Nacional de Avaliação da

Educação Básica – SAEB/1995, SAEB/1997, SAEB/1999 e SAEB/2001. Nesses estudos, pesquisadores fazem uso do modelo logístico de três parâmetros para questões de múltipla escolha, em que se é possível estimar o acerto ao acaso. Para os itens abertos, em que se considera as respostas construídas pelos estudantes, utiliza-se o modelo logístico de dois parâmetros (Klein, 2009).

Por se tratar de um modelo estatístico, por meio da TRI é possível calcular a probabilidade de um sujeito acertar um determinado item e, ainda, estimar a habilidade latente necessária para a resolução de um problema (Hambleton, Swaminathan & Rogers, 1991; Andrade & Valle, 1998; Valle, 2000; Tejada & Meléndez, 2001; Vendramini, 2005; Conde & Laros, 2007; Pasquali & Primi, 2007). 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 teoria do traço latente indica um modelo que se pode expressar, por meio de fórmula matemática, a relação aparente entre variáveis hipotéticas e variáveis observadas, que levam o nome de traços latentes. Os itens de um teste, por exemplo, podem ser considerados variáveis observáveis e, quando se conhece suas características, estas se

tornam constantes na equação e, dessa maneira, solucionáveis. A partir deste 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; Vendramini, 2005; Pasquali & Primi, 2007).

De acordo com Pasquali e Primi (2007), duas características básicas da TRI devem ser consideradas. A primeira indica que as aptidões ou traços latentes (identificados na TRI pela letra grega θ) entendidos como um conjunto de fatores ou variáveis hipotéticas, servem como preditores de desempenho do sujeito em uma determinada tarefa. Tal técnica é conhecida por modelagem latente, em que o θ é a causa e o desempenho o efeito.

A segunda característica menciona a Curva Característica do Item (CCI) que se caracteriza por uma função matemática monotônica crescente, em que se é possível visualizar a relação entre traços latentes e o desempenho do sujeito no item. Sendo assim, a Figura 3 assinala que a medida que a habilidade latente aumenta, proporcionalmente, aumentam as chances, ou seja, a probabilidade de acerto ao item (Pasquali & Primi, 2007).

Deve-se considerar, ao menos, cinco das principais vantagens na utilização da TRI. A primeira delas refere-se ao cálculo do nível de aptidão do sujeito que pode ser realizada independentemente ao número de itens ou conjunto de itens, desde que meçam o mesmo traço latente, isto é, a habilidade do respondente não depende do teste. Outra vantagem diz respeito ao cálculo dos parâmetros dos itens. Isso significa dizer que os cálculos de dificuldade e discriminação de um item independe da amostra utilizada (Hambleton, Swaminathan & Rogers, 1991; Pasquali & Primi, 2007).

A terceira vantagem é que a TRI torna possível o emparelhamento dos itens de acordo com a habilidade do sujeito. Assim, pode-se utilizar itens mais difíceis para sujeitos

com maior habilidade e itens mais fáceis para sujeitos menos aptos, produzindo escores que possam ser comparados nos dois casos. Como quarta vantagem destaca-se a não necessidade de suposições, aparentemente, improváveis como, por exemplo, a premissa de que os erros de medida possam ser iguais para todos os sujeitos. E, por último, diferente das exigências da psicometria clássica, a TRI não apresenta a necessidade de trabalhar com testes exclusivamente paralelos (Hambleton, Swaminathan & Rogers, 1991; Pasquali & Primi, 2007).

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. Destarte, a Teoria de Resposta ao Item se apresenta como uma ferramenta eficaz para a avaliação em larga escala, pois considera a habilidade dos sujeitos em suas respostas à cada item, promovendo bases mais sólidas e precisas para a discussão da questão educacional.

Unidimensionalidade na TRI

A maioria dos dados da TRI assumem que os itens constituem, apenas, uma dimensão, ou seja, são unidimensionais e antes de sua utilização o pesquisador deve assegurar-se que os dados cumprem essa condição (Lord, 1980). Essa é uma restrição

importante para o uso da TRI, já que muito dos dados manejados pelos psicólogos não são, essencialmente, unidimensionais (Muñiz, 2010).

Dessa maneira, os modelos unidimensionais são utilizados quando os dados que serão analisados atendem aos pressupostos especificados como, por exemplo, o critério da independência local, em que as respostas dos sujeitos aos itens não devem ser influenciadas pelas respostas fornecidas a outros itens, quando apenas um único traço é responsável pelo conjunto de itens que compõem a prova; e critério da unidimensionalidade, o qual supõe que os itens do teste devem medir um único traço, ou seja, há apenas um fator dominante influenciando o desempenho dos respondentes do teste (Vendramini, 2005).

Pesquisas como emprego da TRI no Brasil

O campo das pesquisas que empregam os métodos propostos pela TRI segue em grande efervescência, já que o advento de novas tecnologias promove sua utilização segura e confiável. Nesse sentido, Andriola (1998) propôs uma análise comparativa entre a TCT e a TRI, a partir de um estudo de caso, com os dados da Avaliação da Qualidade do Ensino das Escolas Públicas do Estado do Ceará, cuja amostra era composta por 7.576 estudantes provenientes de escolas públicas. Para esse estudo, o autor utilizou-se do modelo logístico de três parâmetros na análise dos itens de um teste de desempenho aplicado aos estudantes e, pôde concluir que esse modelo foi o que permitiu uma análise mais extensa no estudo de casos inicialmente proposto, bem como o melhor modelo que se ajustou à classificação empírica dos resultados.

Também se utilizando do modelo de três parâmetros Pasquali (2005) realizou um estudo cujo objetivo era analisar os parâmetros psicométricos do ENC/Provão de Psicologia, aplicados nos anos de 2000 e 2001. Em 2000 participaram do exame 9.362 estudantes em que, 54,5% eram de universidades particulares. Os resultados indicaram um índice de discriminação (parâmetro a) médio de $a=0,518$ ($DP=0,273$), o que possibilitou afirmar que as questões da prova não apresentaram grande poder discriminativo, haja vista que esse índice pode variar entre 0 (nada discriminativo) até cerca de 4 (extremamente discriminativo). Já o parâmetro b , que refere-se à dificuldade dos itens, apresentou questões com índices muito discrepantes com valores de $b=14,35$ (questão 25) e $b=24,91$ (questão 13). De acordo com os resultados obtidos por esse parâmetro, pôde-se constatar que a prova se mostra muito difícil para a população avaliada. Quanto ao acerto ao acaso (parâmetro c), o resultado foi de 17%, ou seja, abaixo do esperado de 20%. Diante desses índices, o autor pôde concluir que a prova foi mal elaborada e que as questões não representavam a realidade do curso.

Os resultados das análises do ENC/Psicologia 2001 foram semelhantes aos do exame de 2000. Em 2001 participaram do exame 9.201 estudantes, 56,9% da amostra eram estudantes de universidades particulares. A média do índice de discriminação dos itens foi $a=0,510$ ($DP=0,227$), o que demonstra a precariedade do poder discriminativo dos itens. O parâmetro b teve como média o valor de $b=0,445$ e também apresentou, como o exame do ano anterior, índices discrepantes (questão 15 com $b=24,97$ e questão 23 com $b=8,27$). Já o índice do parâmetro c , igualmente, foi abaixo do esperado, atingindo o valor de 14%. Conclui-se que, embora semelhantes, os resultados demonstraram 37 itens em melhores condições psicométricas em relação ao Provão de 2000. No entanto, a prova apresentou

falhas em sua construção, principalmente, no que diz respeito ao conhecimento do caráter epistemológico e de fundamentos em Psicologia (Pasquali, 2005).

Ainda no contexto da avaliação do ensino superior, o estudo de Oliveira (2006) objetivou a análise das propriedades psicométricas do ENADE por meio da TRI. A autora utilizou-se de um banco de dados contendo 8.124 respostas dos estudantes de Medicina a 28 questões objetivas das provas de Formação Geral e Componente Específico do ENADE, aplicado no ano de 2004. Com a utilização das ferramentas fornecidas pela TRI, a pesquisadora averiguou que os dados de seu estudo se ajustaram com maior confiabilidade ao modelo logístico de um parâmetro, ou modelo de Rasch, nos dois componentes do ENADE. Pôde-se concluir que, nos modelos de 2 e 3 parâmetros, não foram atendidas os critérios mínimos para o emprego da TRI.

Em outro estudo semelhante, Nogueira (2006) propôs aplicar a TRI na avaliação das questões da prova de Formação Geral do ENADE, em especial aquelas que envolviam conceitos estatísticos, visando o ajuste dos itens ao modelo de Rasch. A autora utilizou dois bancos de dados contendo informações acadêmicas de 403.512 estudantes submetidos ao ENADE nos anos de 2004 e 2005. Os resultados indicaram que as provas se ajustam ao modelo de Rasch devido, principalmente, a baixa frequência de resultados inesperados de padrões de erros e acertos.

Com o objetivo de avaliar a prova de Formação Geral do ENADE por meio da TRI, Lopes (2009) utilizou-se de um banco de dados com as respostas de 49.497 estudantes do curso de Pedagogia que participaram do ENADE no ano de 2005. O modelo empregado na análise dos dados foi o modelo de um parâmetro (Rasch), que permitiu indicar os itens que

exigiam maior ou menor habilidade dos participantes do exame. Também para esse estudo, o modelo de Rasch foi o que mais se ajustou aos padrões de resposta.

Exemplos do uso da TRI na avaliação do Ensino Superior são propostos por Primi e cols. (2009), Vendramini (2009) e Primi e cols. (2010). O estudo sugerido por Primi e cols. (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, uma das ferramentas da TRI. Para tanto, os pesquisadores utilizaram um banco de dados contendo as respostas ao ENADE de 26.613 estudantes do curso de Psicologia de todas as regiões do país. 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 *theta* e, por fim, transformada em uma nota de 0 a 100, de acordo com a porcentagem de acertos dos estudantes.

Os resultados encontrados pelos pesquisadores indicaram que as medidas de dificuldade dos itens e das habilidades dos estudantes foram adequadas, não existindo a ocorrência de padrões de respostas inesperados. Para estabelecer os pontos de corte foi utilizado o mapa de itens que possibilitou a distinção dos itens em 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, estão os 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 & cols., 2009).

Vendramini (2009) buscou verificar 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 as 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, Biblioteconomia, 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 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 (Vendramini, 2009).

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).

Primi e cols. (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 26613 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 alunos ingressantes e concluintes, cinco itens não atingiram o critério mínimo de DIF, sendo que três itens favoreceram os ingressantes e dois itens favoreceram os alunos concluintes. Já para os estudantes de universidades públicas e privadas, sete itens atingiram o critério mínimo de DIF, em que 6 itens favoreceram alunos de instituições públicas.

Segundo os autores, um grupo menor de itens com DIF indicaram problemas em sua formulação, favorecendo os alunos 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, a associação desses itens ser maior no que se refere aos sistemas de ensino do que aos momentos de formação do aluno.

Considerando os resultados das pesquisas acima citadas, para este estudo, o modelo utilizado para a análise dos dados é o modelo logístico de um parâmetro, ou 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, permitindo uma aplicação prática na estimativa de parâmetros (Perline, Wright & Wainer, 1979).

Equalização e TRI

A equalização, na Teoria de Resposta ao Item, 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 habilidades, já que essa ferramenta possibilita que os parâmetros de itens e/ou de habilidades dos respondentes, de grupos diferentes, possam ser analisados e compreendidas 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, equalização via itens comuns em que calibram-se 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 na 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, informações que mencionem a capacidade do estudante necessária a 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 PAC.

Pesquisas que fazem uso da equalização na situação de Avaliação da Educação Superior ainda são escassas. Nesse sentido, destaca-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. (2008a; 2008b; 2008c;

2009a; 2009b; 2009c; 2009d; 2009e; 2009f; 2009g; 2009h; 2009i; 2010a; 2010b; 2010c; 2010d; 2010e; 2010f; 2010g; 2010h; 2010i; 2010j; 2010k), 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 propuseram-se os objetivos descritos a seguir.

OBJETIVOS

Objetivo Geral

Equalizar uma prova de habilidades acadêmicas, de um Programa de Avaliação Continuada – PAC em Pedagogia, com o Exame Nacional de Desempenho dos Estudantes – ENADE/2005, por meio da Teoria de Resposta ao Item.

Objetivos específicos

- Analisar as propriedades psicométricas dos itens que compõem a prova de Pedagogia do ENADE 2005;
- Selecionar itens que apresentem as melhores propriedades psicométricas para compor a prova PAC de Pedagogia;
- Analisar e interpretar os mapas de itens do ENADE de Pedagogia de 2005 e do PAC de Pedagogia, obtidos pela TRI;
- Verificar as possíveis diferenças no desempenho acadêmico dos estudantes que fizeram uma prova PAC de Pedagogia e os que fizeram o ENADE 2005 de Pedagogia.

CAPÍTULO 4

RESULTADOS E DISCUSSÃO

Com o intento de atender os objetivos explicitados, para esta pesquisa propôs-se três estudos, um primeiro para avaliar e selecionar os itens do ENADE 2005 de Pedagogia; um segundo para analisar uma prova PAC de Pedagogia equalizada ao ENADE 2005; e, um terceiro estudo para analisar e interpretar os resultados de um grupo de estudantes que fizeram o PAC comparados com todos os estudantes brasileiros. Cada estudo é apresentado separadamente e são descritos a seguir.

4.1 *Estudo 1* – Análise e Seleção dos Itens do ENADE 2005 de Pedagogia

4.1.1 Método

Fonte de dados

Os participantes desta etapa do estudo 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 monitorado pelo Ministério da Educação. 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 Sócio-econômico.

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%). Na Tabela 4 é possível visualizar a distribuição dos estudantes de Pedagogia de acordo com a categoria administrativa das IES em que, a maior parte (70,7%) dos participantes está matriculada em Instituições de Ensino Superior privadas.

Tabela 4

Distribuição de estudantes no ENADE 2005 de Pedagogia por categoria administrativa

Categoria administrativa	Número de estudantes	%
Federal	5.346	10,8
Estadual	7.920	16,0
Municipal	1.230	02,5
Particular	35.001	70,7
Total	49.497	100,0

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 à parte do 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, aprovadas pelo Conselho Nacional de Educação e no perfil profissional de cada carreira, levando em consideração os saberes fundamentais em cada área profissional (MEC, 2005). O estudo 1 foi realizado em duas etapas descritas a seguir.

Etapa 1 – Análise das provas de Pedagogia do ENADE via TRI

Procedimento

Após a aprovação do Comitê de Ética em Pesquisa da Universidade São Francisco para a realização da pesquisa, os dados foram organizados a fim de se atender os pressupostos do método estatístico proposto para o estudo. Ressalta-se que cada etapa deste estudo contou com procedimentos específicos.

Etapa 2 – Seleção de itens para compor uma prova PAC

Procedimento de análise dos dados

Foram extraídas da prova do ENADE as questões que compuseram a prova do PAC/Pedagogia cujos critérios são embasados na Teoria de Resposta ao Item. Assim, o primeiro critério diz respeito à Correlação item total, em que são considerados ideais itens com valores acima de 0,30; pode-se aceitar a escolha de alguns itens a partir de 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 nessas duas análises valores variando entre 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. Para compor uma prova de Pedagogia do PAC, foram selecionados itens fáceis, medianos e

difíceis, considerando as informações apresentadas na Tabela 5. E, por último, a questão das cargas dos resíduos completam 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 seleção dos itens considera que o *residual loading* seja menor que 0,40. Cabe ainda destacar que será considerada como ideal, a seleção de itens que contemplem no mínimo três dos critérios expostos (Linacre & Wright, 1991).

Tabela 5

Níveis de dificuldade dos itens quando estimados por TRI

Valor de “b”	Faixa
$-4 < b \leq -2$	Muito Fácil
$-2 < b \leq -1$	Fácil
$-1 < b \leq +1$	Mediano
$+1 < b \leq +2$	Difícil
$+2 < b \leq +4$	Muito Difícil

Foi solicitado à coordenação do curso de Pedagogia que, com base nos parâmetros psicométricos dos itens do ENADE, selecionassem aqueles que satisfizessem os critérios estabelecidos para este estudo. Após a seleção dos itens que compuseram a prova de Pedagogia do PAC, foi realizada a aplicação da mesma de forma coletiva em sala de aula, em dias e horários pré-estabelecidos pela IES. Os aplicadores foram os docentes selecionados previamente pelo colegiado do curso de Pedagogia.

4.1.2 Resultados e Discussão

Etapa 1 – Análise das provas de Pedagogia do ENADE via TRI

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 (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 (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 Diretrizes Curriculares Nacionais 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 à conhecimentos pedagógicos de formação geral, conhecimentos pedagógicos-didáticos e conhecimentos das áreas específicas. A Tabela 6 abarca os conteúdos solicitados aos estudantes nesse componente (Inep, 2005).

Tabela 6

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: Inep (2005) – Relatório Síntese

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 7 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 7

Estatísticas básicas das questões objetivas de Formação Geral por grupo de estudantes

Estatísticas	Total	Grupo	
		Ingressantes	Concluintes
População	127,153	59,257	67,896
Tamanho da amostra	55,240	27,188	28,052
Presentes	49,497	23,318	26,179
Média	63,6	62,5	64,6
Erro padrão da média	0,1	0,1	0,1
Desvio-padrão	22,4	22,5	22,2
Nota mínima	0,0	0,0	0,0
Mediana	71,4	71,4	71,4
Nota máxima	100,0	100,0	100,0

Fonte: Inep (2005) – Relatório Síntese

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 entre 85% e 61%; do mesmo modo, para serem considerados itens de dificuldade mediana, as questões deveriam obter um índice de acerto

entre 60% e 41%; já para 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 8 são apresentados os itens correspondentes as 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 (Inep, 2005).

Tabela 8

Classificação das questões objetivas de Formação Geral por índice de facilidade

Índice de Facilidade (%)	Classificação	Itens
≥ 86	Muito fácil	
61 a 85	Fácil	1, 2, 3, 4, 7
41 a 60	Mediana	5
16 a 40	Difícil	6
≤ 15	Muito Difícil	

Fonte: Inep (2005) – Relatório Síntese

Para o componente de Conhecimento Específico foram realizadas as mesmas análises em que se obtiveram as estatísticas apresentadas na Tabela 9. 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 foi 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 9

Estatísticas básicas das questões objetivas de Conhecimento Específico por grupo de estudantes

Estatísticas	Total	Grupo	
		Ingressantes	Concluintes
População	127,153	59,257	67,896
Tamanho da amostra	55,240	27,188	28,052
Presentes	49,497	23,318	26,179
Média	45,7	41,5	49,4
Erro padrão da média	0,1	0,1	0,1
Desvio-padrão	16,9	16,0	16,7
Nota mínima	0,0	0,0	0,0
Mediana	46,2	42,3	50,0
Nota máxima	96,2	96,2	92,3

Fonte: Inep (2005) – Relatório Síntese

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 10, 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 (Inep, 2005).

Tabela 10

Classificação das questões objetivas do Componente Específico por índice de facilidade

Índice de Facilidade (%)	Classificação	Itens
≥ 86	Muito fácil	17
61 a 85	Fácil	16, 21, 22, 32
41 a 60	Mediana	8, 12, 15, 19, 23, 25, 27, 30, 31, 33, 34
16 a 40	Difícil	9, 11, 14, 18, 20, 24, 26, 28, 29
≤ 15	Muito Difícil	13

Fonte: Inep (2005) – Relatório Síntese

Esses resultados corroboram as estimativas da qualidade dos cursos avaliados pelo ENADE. Sendo assim, por meio da medição das habilidades 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. A Tabela 11 contém os cálculos finais das notas dos estudantes e seus respectivos conceitos.

Tabela 11

Distribuição dos conceitos

Conceito	Notas finas
1	0,0 a 0,9
2	1,0 a 1,9
3	2,0 a 2,9
4	3,0 a 3,9
5	4,0 a 5,0

Fonte: Inep (2005) – Relatório Síntese

De acordo com os critérios expostos na Tabela 11 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 12, que a maior parte dos cursos obtiveram o conceito 3, o que equivale a 41,9% dos cursos participantes. A Tabela 12 também traz informações sobre as regiões do país e a respectiva distribuição dos conceitos em que pode-se observar que a região Sudeste apresentou mais cursos avaliados com o conceito 3 (Inep, 2005).

Tabela 12

Distribuição das regiões do Brasil por conceito do curso no ENADE de Pedagogia de 2005

Conceito	Brasil		Região									
			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: Inep (2005) – Relatório Síntese

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 (Linacre & 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 13. De acordo com essa tabela, a questão que solicitou menor habilidade dos respondentes refere-se ao item 4 ($b=-1,43$) que, dentre os critérios estabelecidos para esta etapa, pode ser considerado um item 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 (Inep, 2005).

Tabela 13

Estatísticas dos itens em Formação Geral no ENADE de Pedagogia 2005 calibradas pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	<i>Residual Loading</i>
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 esta etapa do estudo, os índices gerados por meio da TRI podem ser considerados suficientes, 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 medida *Residual Loading*, duas questões 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 distribuídos na Tabela 14 em que se pode observar índices de

dificuldade que indicam itens muito fáceis, fáceis, medianos, difíceis e muito difíceis. Com base nesses resultados e com 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 o item 17 obteve uma porcentagem de acerto maior ou igual a 86%, denominado como muito fácil e, o item de número 13, que obteve a menor porcentagem de acerto ($\leq 15\%$), avaliado como muito difícil (Inep, 2005).

Conforme os dados referentes à Tabela 14 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. Já para os valores da medida entendida como Correlação item total, os itens de número 11 ($r_{item-total} = 0,21$), 12 ($r_{item-total} = 0,28$), 13 ($r_{item-total} = 0,20$), 24 ($r_{item-total} = 0,24$), 25 ($r_{item-total} = 0,25$), 26 ($r_{item-total} = 0,23$), 29 ($r_{item-total} = 0,18$) e 31 ($r_{item-total} = 0,28$), tiveram suas correlações abaixo de 0,30, valor entendido como ideal, totalizando 30,77% dos itens. Cabe ressaltar que pode-se aceitar como suficiente itens com medidas a partir de 0,20, se obedecerem aos demais critérios, bem como a relevância do conteúdo avaliado.

Tabela 14

Estatísticas dos itens no Componente Específico do ENADE de Pedagogia 2005 calibradas

Item	Escore Bruto	N	Dificuldade	Infit	Outfit	Correlação Item-total	Residual Loading
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		

A Figura 3 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/dificuldade dos itens, 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.

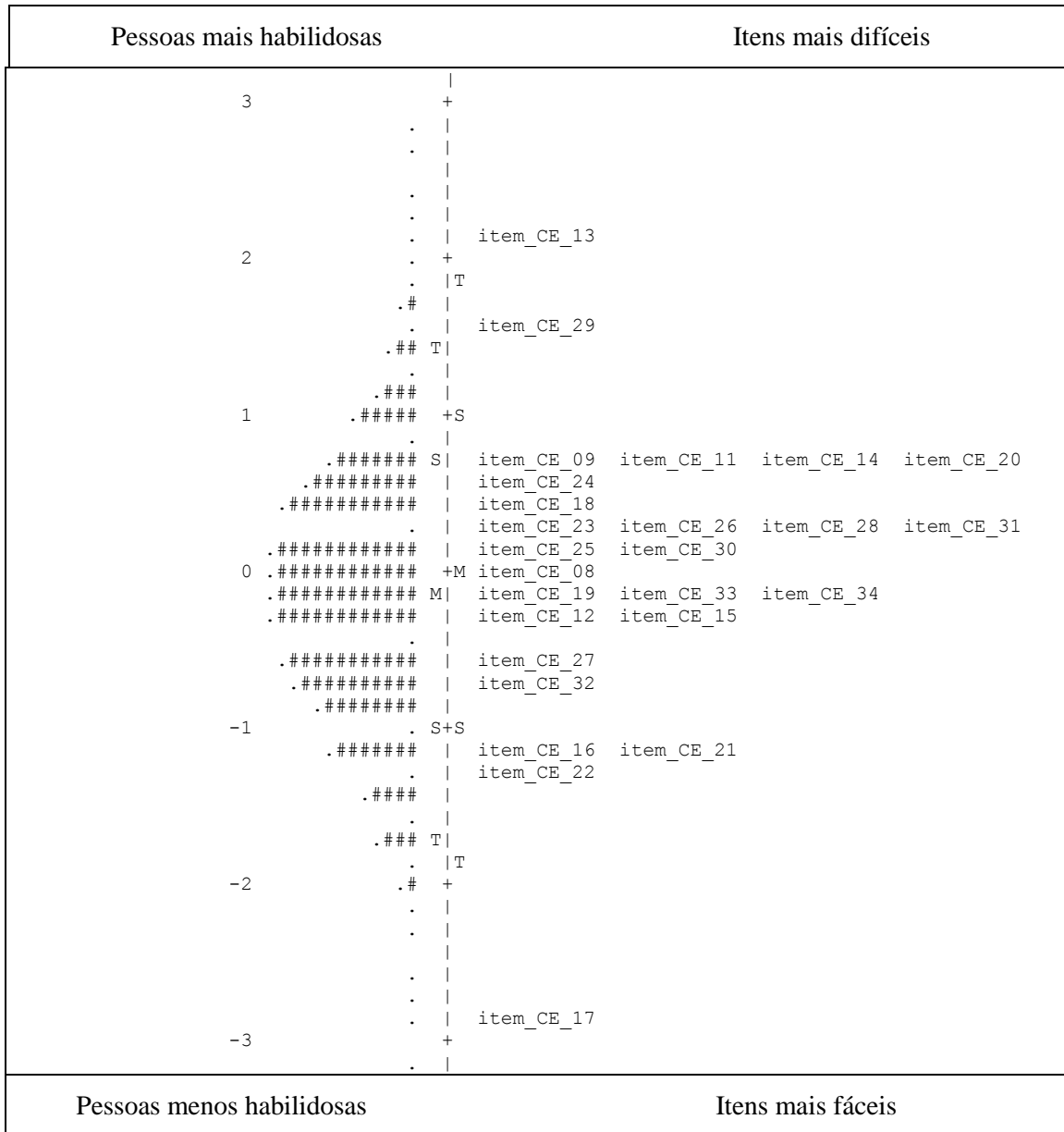


Figura 3. Comparação de habilidades dos estudantes e dificuldade dos itens

A Tabela 15 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. Por esses dados, identifica-se que, em média, os itens apresentaram dificuldade mediana e houve indicadores que alcançaram até 2,17 de dificuldade (máximo) e -2,91 (mínimo). Também, no geral, os índices de ajuste de *infit* e *outfit* podem ser considerados suficientes. 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).

Tabela 15

Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas e itens.

Parâmetro dos itens

Estatísticas	Score Bruto	Dificuldade	Infit	Outfit
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

Parâmetros das pessoas

Estatísticas	Score Bruto	Dificuldade	Infit	Outfit
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

Etapa 2 – Seleção dos itens da prova PAC

Para esta etapa, seguiu-se os critérios estabelecidos e chegou-se à seguinte escolha dos itens. Os itens do componente de Formação Geral que serviriam de itens âncora, ou seja, itens que fazem parte do ENADE 2005 e comuns ao PAC, deveriam perfazer o total de 5 questões. Ressalta-se que as outras 5 questões que compuseram o componente de Formação Geral do PAC, foram escolhidas pelos colegiados dos cursos. Dessa maneira, foram selecionados os itens descritos abaixo com seus respectivos gabaritos.

Leia e relacione os textos a seguir.

O Governo Federal deve promover a inclusão digital, pois a falta de acesso às tecnologias digitais acaba por excluir socialmente o cidadão, em especial a juventude.

(Projeto Casa Brasil de inclusão digital começa em 2004.

In: MAZZA, Mariana. JB online.)

Comparando a proposta acima com a charge, pode-se concluir que

- (A) o conhecimento da tecnologia digital está democratizado no Brasil.
- (B) a preocupação social é preparar quadros para o domínio da informática.
- (C) o apelo à inclusão digital atrai os jovens para o universo da computação.
- (D) o acesso à tecnologia digital está perdido para as comunidades carentes.
- (E) a dificuldade de acesso ao mundo digital torna o cidadão um excluído social.



Fonte: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos-2005>

Figura 4. Questão de número 1 do PAC e correspondente à questão 2 do componente de Formação Geral do ENADE 2005.

A habilidade aferida nesta questão (Figura 4), de acordo com o Inep (2005), é de estabelecer relações entre gêneros textuais distintos, mas com a mesma temática que, neste caso, refere-se à inclusão/exclusão digital e a alternativa correta é a E. Os parâmetros da TRI para essa questão estão de acordo com os critérios já estabelecidos em que a correlação item-total foi de 0,46, as medidas de *Infit* e *Outfit* também se apresentaram suficientes com 0,97 e 0,96, respectivamente. O índice de dificuldade do item considerou a questão como de dificuldade mediana ($b = -0,45$) e, o critério referente à medida *Residual Loading* manteve-se adequado, -0,16. Então, para este item, todos os critérios foram atendidos.

As ações terroristas cada vez mais se propagam pelo mundo, havendo ataques em várias cidades, em todos os continentes. Nesse contexto, analise a seguinte notícia:

No dia 10 de março de 2005, o Presidente de Governo da Espanha José Luis Rodriguez Zapatero em conferência sobre o terrorismo, ocorrida em Madri para lembrar os atentados do dia 11 de março de 2004, “assinalou que os espanhóis encheram as ruas em sinal de dor e solidariedade e dois dias depois encheram as urnas, mostrando assim o único caminho para derrotar o terrorismo: a democracia. Também proclamou que não existe alibi para o assassinato indiscriminado. Zapatero afirmou que não há política, nem ideologia, resistência ou luta no terror, só há o vazio da futilidade, a infâmia e a barbárie. Também defendeu a comunidade islâmica, lembrando que não se deve vincular esse fenômeno com nenhuma civilização, cultura ou religião. Por esse motivo apostou na criação pelas Nações Unidas de uma aliança de civilizações para que não se continue ignorando a pobreza extrema, a exclusão social ou os Estados falidos, que constituem, segundo ele, um terreno fértil para o terrorismo”.

(MANCEBO, Isabel. Madri fecha conferência sobre terrorismo e relembra os mortos de 11-M. (Adaptado).

A principal razão, indicada pelo governante espanhol, para que haja tais iniciativas do terror está explicitada na seguinte afirmação:

- (A) O desejo de vingança desencadeia atos de barbárie dos terroristas.
- (B) A democracia permite que as organizações terroristas se desenvolvam.
- ➔ (C) A desigualdade social existente em alguns países alimenta o terrorismo.
- (D) O choque de civilizações aprofunda os abismos culturais entre os países.
- (E) A intolerância gera medo e insegurança criando condições para o terrorismo.

Fonte: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos-2005>

Figura 5. Questão de número 2 do PAC e correspondente à questão 3 do componente de Formação Geral do ENADE 2005.

A questão 3 do componente de Formação Geral do ENADE 2005 (Figura 5), selecionada para ser parte do PAC, aferiu a habilidade de identificar e analisar a ideia central de textos e apresentou os parâmetros obtidos pela TRI considerados suficientes (Inep, 2005). Assim, o primeiro critério que diz respeito à correlação item-total foi de 0,50, a medida de *Infit* alcançou o valor de 0,95, o *Outfit* foi de 0,92 e, o *Residual Loading* dessa questão foi de -0,15. A dificuldade do item foi considerada mediana ($b = -0,07$) e, mais uma vez, todos os critérios propostos foram apreciados.

(Laerte. O condomínio)

(Laerte. O condomínio)

(Disponível em: <http://www2.uol.com.br/laerte/tiras/index-condomínio.html>)

As duas charges de Laerte são críticas a dois problemas atuais da sociedade brasileira, que podem ser identificados pela crise

➔ (A) na saúde e na segurança pública.
 (B) na assistência social e na habitação.
 (C) na educação básica e na comunicação.
 (D) na previdência social e pelo desemprego.
 (E) nos hospitais e pelas epidemias urbanas.

Fonte: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos-2005>

Figura 6. Questão de número 3 do PAC e correspondente à questão 4 do componente de Formação Geral do ENADE 2005.

Para essa questão do ENADE (Figura 6) pretendeu-se aferir a habilidade de identificar e criticar os problemas sociais vigentes na ocasião da prova (Inep, 2005). As medidas obtidas por meio da TRI foram consideradas adequadas e suficientes para atestar a qualidade do item e a habilidade dos respondentes em que a correlação item-total foi de 0,38, a medida de *Infit* foi de 0,96 e a de *Outfit*, foi de 0,83. De acordo com o índice de dificuldade do item, a questão foi considerada a mais fácil de todo o componente de Formação Geral ($b = -1,62$); a medida de *Residual Loading* foi de -0,16.

Leia trechos da carta-resposta de um cacique indígena à sugestão, feita pelo Governo do Estado da Virgínia (EUA), de que uma tribo de índios enviasse alguns jovens para estudar nas escolas dos brancos.

“(…) Nós estamos convencidos, portanto, de que os senhores desejam o nosso bem e agradecemos de todo o coração. Mas aqueles que são sábios reconhecem que diferentes nações têm concepções diferentes das coisas e, sendo assim, os senhores não ficarão ofendidos ao saber que a vossa idéia de educação não é a mesma que a nossa. (...) Muitos dos nossos bravos guerreiros foram formados nas escolas do Norte e aprenderam toda a vossa ciência. Mas, quando eles voltaram para nós, eram maus corredores, ignorantes da vida da floresta e incapazes de suportar o frio e a fome. Não sabiam caçar o veado, matar o inimigo ou construir uma cabana e falavam nossa língua muito mal. Eles eram, portanto, inúteis. (...) Ficamos extremamente agradecidos pela vossa oferta e, embora não possamos aceitá-la, para mostrar a nossa gratidão concordamos que os nobres senhores de Virgínia nos enviem alguns de seus jovens, que lhes ensinaremos tudo que sabemos e faremos deles homens.”

(BRANDÃO, Carlos Rodrigues. O que é educação. São Paulo: Brasiliense, 1984)

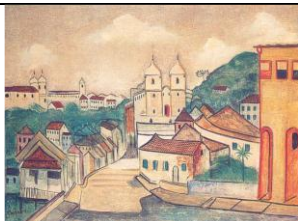
A relação entre os dois principais temas do texto da carta e a forma de abordagem da educação privilegiada pelo cacique está representada por:

- (A) sabedoria e política / educação difusa.
- (B) identidade e história / educação formal.
- (C) ideologia e filosofia / educação superior.
- (D) ciência e escolaridade / educação técnica.
- ➔ (E) educação e cultura / educação assistemática.

Fonte: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos-2005>

Figura 7. Questão de número 4 do PAC e correspondente à questão 5 do componente de Formação Geral do ENADE 2005.

Nesta questão (Figura 7) buscou-se investigar a habilidade na análise das diferentes abordagens do processo educacional em culturas distintas, bem como, a habilidade de fazer comparações e detectar contradições nesses tipos de abordagens (Inep, 2005). Este item pôde ser considerado difícil, de acordo com seu índice de dificuldade que foi de 1,17. Os demais parâmetros, também considerados suficientes, detectaram o valor de 0,47 em sua correlação item-total, 1,07 e 1,12 nas medidas de *Infit* e *Outfit*, respectivamente. Já a medida de *Residual Loading* obteve o valor de 0,97 o que, de acordo com os critérios estabelecidos para esta medida, não pode ser considerada suficiente; contudo, o item foi adequado e suficiente para três dos quatro critérios, que justifica sua escolha.



(Coleção Roberto Marinho. Seis décadas da arte moderna brasileira. Lisboa: Fundação Calouste Gulbenkian, 1989. p.53.)

A “cidade” retratada na pintura de Alberto da Veiga Guignard está tematizada nos versos

(A) Por entre o Beberibe, e o oceano
Em uma areia sáfia, e lagadiça
Jaz o Recife povoação mestiça,
Que o belga edificou ímpio tirano.

(MATOS, Gregório de. Obra poética. Ed. James Amado. Rio de Janeiro: Record, 1990. Vol. II, p. 1191.)

➔ (B) Repousemos na pedra de Ouro Preto,
Repousemos no centro de Ouro Preto:
São Francisco de Assis! igreja ilustre, acolhe,
À tua sombra irmã, meus membros lassos.

(MENDES, Murilo. Poesia completa e prosa. Org. Luciana Stegagno Picchio. Rio de Janeiro: Nova Aguilar, 1994. p. 460.)

(C) Bembelelém
Viva Belém!
Belém do Pará porto moderno integrado na equatorial
Beleza eterna da paisagem
Bembelelém
Viva Belém!

(BANDEIRA, Manuel. Poesia e prosa. Rio de Janeiro: Aguilar, 1958. Vol. I, p. 196.)

(D) Bahia, ao invés de arranha-céus, cruzeiros e cruzeiros
De braços estendidos para os céus,
E na entrada do porto,
Antes do Farol da Barra,

O primeiro Cristo Redentor do Brasil!

(LIMA, Jorge de. Poesia completa. Org. Alexei Bueno. Rio de Janeiro: Nova Aguilar, 1997. p. 211.)

(E) No cimento de Brasília se resguardam
maneiras de casa antiga de fazenda,
de copiar, de casa-grande de engenho,
enfim, das casarões de alma fêmea.

(MELO NETO, João Cabral. Obra completa. Rio de Janeiro: Nova Aguilar, 1994. p. 343.)

Fonte: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos-2005>

Figura 8. Questão de número 5 do PAC e correspondente à questão 7 do componente de Formação Geral do ENADE 2005.

Por fim, o item 7 (Figura 8) teve por objetivo aferir a habilidade de reconhecimento de diferentes tipos de manifestações artísticas tais como a pintura, a poesia, a literatura e a arquitetura em relação ao patrimônio cultural brasileiro (Inep, 2005). A análise deste item por meio da TRI indicou os seguintes resultados, para a correlação item-total o valor foi de 0,42; para as medidas de *Infit* e *Outfit* os valores respectivos foram 1,04 e 1,04 e, na medida de *Residual Loading* o valor obtido foi de -0,03; o item foi considerado de dificuldade mediana ($b = 0,42$). Para este item, todos os critérios propostos para essa etapa do estudo foram contemplados.

Escolha dos itens do Componente Específico

A escolha dos itens do Componente Específico que compuseram o PAC de Pedagogia seguiu os mesmos requisitos utilizados na escolha dos itens de Formação Geral, acima descritos. Ressalta-se que foram indicados 11 itens para apreciação do colegiado do curso de Pedagogia com relação à forma e ao conteúdo da questão e, após análise dos coordenadores e professores do curso, os itens foram mencionados como pertinentes e suficientes para os propósitos da avaliação. Esse mesmo colegiado selecionou outras 19 questões referentes ao ENADE aplicado ao curso de Pedagogia no ano de 2008 para completar a prova. Contudo, para este estudo, os itens de 2008 não foram considerados como itens âncoras.

A fim de melhor ilustração dos resultados obtidos pela TRI, as tabelas abaixo tiveram por objetivo trazer as questões selecionadas para compor o PAC de Pedagogia, bem como divulgar informações acerca dos conteúdos predominantes e as habilidades aferidas para cada questão. Essas informações são parte de um Relatório Síntese da área de

Pedagogia do ENADE 2005, elaboradas de acordo com os profissionais que atuam junto ao INEP/MEC e disponíveis para consulta pública na rede mundial de computadores.

<p>A primeira iniciativa, no Brasil, de formar o professor primário em nível universitário se deu com a Escola de Educação da Universidade do Distrito Federal, em 1935. Liderada pelo intelectual Anísio Teixeira, esta iniciativa contribuiu na configuração do perfil e da carreira do educador, com a definição de um espaço de atuação profissional precisamente identificado. Contudo, para que esta formação docente obtivesse resultados positivos, era preciso</p> <p>(A) conceber a atividade educativa somente como prática ascética.</p> <p>(B) preparar o professor para assumir funções técnicas e administrativas.</p> <p>→(C) definir a educação como arte prática e instrumento de análise das Ciências Sociais.</p> <p>(D) construir conhecimentos educacionais independentes das demais Ciências Humanas.</p> <p>(E) dispensar o exercício científico na ação educativa a partir da perspectiva da organização da escola.</p>	
Conteúdo	Habilidades
História da Educação/ História da Educação Brasileira/ História da Pedagogia e do Curso de Pedagogia	Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais.

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 51.

Figura 9. Questão Correspondente ao item 9 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

A análise deste item (Figura 9) pela TRI atingiu os seguintes resultados, para a correlação item-total o valor foi de 0,32, considerado adequado e suficiente; a medida de *Infit* obteve o valor de 1,02 e, a de *Outfit*, o valor de 1,03 – medidas também consideradas suficientes e adequadas de acordo com os critérios já estabelecidos. O índice de dificuldade indicou que este item apresentou dificuldade mediana ($b=0,70$) e, a medida referente ao *Residual Loading* atingiu o valor de 0,03 – valores consonantes com a qualidade pretendida dos itens.

<p>Uma teoria pedagógica é um conjunto de saberes sobre as questões principais da Pedagogia: Para que educar? O que significa ensinar e aprender? Como fazê-lo? Uma teoria pedagógica crítica se caracteriza pela</p> <p>(A) visão pessimista da escola e do papel sociocultural que ela desempenha na sociedade.</p> <p>→(B) problematização dos pressupostos filosóficos e sociopolíticos do fazer pedagógico.</p> <p>(C) instrumentalização eficiente do ensinar e do aprender.</p> <p>(D) rejeição aos métodos didáticos que simplificam a capacidade de alunos e de professores.</p> <p>(E) aceitação de diferentes propostas sobre o significado do ensinar e do aprender.</p>	
Conteúdo	Habilidades
Teorias Pedagógicas	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Articular as teorias pedagógicas e de currículo no desenvolvimento da docência, na elaboração e avaliação de projetos pedagógicos, na organização e gestão do trabalho educativo escolar e não escolar.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 51.

Figura 10. Questão Correspondente ao item 14 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

Para este item (Figura 10) foi observado os seguintes valores obtidos por meio da TRI, no primeiro critério que trata da correlação item-total, o valor foi de 0,39 – maior que o ponto crítico de 0,30; para as medidas de *Infiti* e *Outfit* os valores foram, respectivamente, 0,94 e 0,93 – compatíveis com os valores pré-estabelecidos; para a medida de *Residual Loading*, que deveria ficar abaixo do valor crítico de 0,40, o valor correspondente a este item foi 0,32. A dificuldade desta questão pode ser considerada como de dificuldade média dado seu parâmetro $b=0,70$. Assim, para esta questão, todos os critérios foram privilegiados.

<p>Durante todo o curso de Pedagogia, Júlia teve a oportunidade de conhecer as diferentes tendências, concepções ou teorias educacionais que dão suporte à organização do trabalho da escola e à prática educativa. Com isso, percebeu a necessidade de se contrapor à fragmentação, à rotina, ao autoritarismo e à centralização do poder. Um dos caminhos a serem trilhados é a construção do projeto político-pedagógico como um instrumento de luta em busca da qualidade e da almejada cidadania. Qual é a ação fundamental para que uma escola tenha o projeto político-pedagógico pretendido?</p> <p>(A) Planejamento de metas pela equipe gestora para os segmentos e áreas do conhecimento.</p> <p>→ (B) Diagnóstico permanente da realidade escolar com registro dos dados e das discussões.</p> <p>(C) Elaboração do projeto por uma coordenação administrativo-pedagógica especializada.</p> <p>(D) Utilização de técnica e de metodologias específicas para a elaboração de projetos.</p> <p>(E) Divulgação de dados de desempenho para uma classificação das turmas.</p>	
Conteúdo	Habilidades
Coordenação, elaboração e avaliação do projeto pedagógico.	<p>Participar da formulação, implementação e avaliação contínua de projetos pedagógicos escolares e não escolares;</p> <p>Planejar, desenvolver e avaliar situações de ensino e de aprendizagem, de modo a adequar objetivos, conteúdos e metodologias específicos das diferentes áreas à diversidade dos alunos e aos fins da educação;</p> <p>Promover e planejar ações visando a gestão democrática nos espaços e sistemas escolares e não escolares.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 51.

Figura 11. Questão Correspondente ao item 15 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

De acordo com as análises resultantes da aplicação da TRI, este item (Figura 11) foi o único que não cumpriu todos os critérios estabelecidos dada sua medida de *Residual Loading* que obteve o valor de 0,47, ou seja, acima do ponto crítico de 0,40. Contudo, as demais medidas atingiram os valores suficientes como se segue, correlação item-total de 0,46, *Infit* de 0,90 e *Outfit* de 0,89, o parâmetro *b* que indica a dificuldade do item foi de -0,28, classificando-o como de dificuldade média.

A notícia veio de supetão: iam meter-me na escola. Já me haviam falado nisso, em horas de zanga, mas nunca me convencera que realizassem a ameaça. A escola, segundo informações dignas de crédito, era um lugar para onde se enviavam as crianças rebeldes. Eu me comportava direito: encolhido e morno, deslizava como sombra. [...] A escola era horrível – e eu não podia negá-la, como negara o inferno. Considerei a resolução de meus pais uma injustiça. [...] Lembrei-me do professor público, austero e cabeludo, arrepiei-me calculando o rigor daqueles braços. Não me defendi, não mostrei as razões que me fervilhavam na cabeça, a mágoa que me inchava o coração. Inútil qualquer resistência. (RAMOS, Graciliano. Infância, Rio de Janeiro: Record, 1995, p. 104.) O texto do escritor Graciliano Ramos traz lembranças de sua entrada na escola, que expressam um momento da Educação brasileira. Entretanto, o pensamento pedagógico tem-se modificado ao longo do tempo, contrapondo-se ao modelo de escola evidenciado no texto. Este contraponto é expresso por:

I - transmissão cultural que considera o aluno como um ser passivo, atribuindo caráter dogmático aos conteúdos de ensino;

II - valorização da criança, do afeto entre professor e aluno, das reflexões sobre as formas de ensino que considerem o saber das crianças;

III - dimensão dialógica do processo ensino/aprendizagem com ênfase nas relações igualitárias;

IV - preocupação com a formação humana relacionando as dimensões humanas, econômica, social, política e cultural.

São corretos:

(A) I e II, apenas.

(B) III e IV, apenas.

(C) I, II e III, apenas.

→ (D) II, III e IV, apenas.

(E) I, II, III e IV.

Conteúdo	Habilidades
Didática	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Compreender o desenvolvimento e a aprendizagem de crianças, jovens e adultos, considerando as dimensões cognitivas, afetivas, socioculturais, éticas e estéticas;</p> <p>Articular as teorias pedagógicas e de currículo no desenvolvimento da docência, na elaboração e avaliação de projetos pedagógicos, na organização e gestão do trabalho educativo escolar e não escolar;</p> <p>Planejar, desenvolver e avaliar situações de ensino e de aprendizagem, de modo a adequar objetivos, conteúdos e metodologias específicos das diferentes áreas à diversidade dos alunos aos fins da educação.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 52.

Figura 12. Questão Correspondente ao item 16 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.


O item acima descrito (Figura 12) exigia, de acordo com o Inep (2005), maior habilidade no reconhecimento dos assuntos tratados na questão. No entanto, conforme os resultados indicaram, pode ser considerado um item fácil, dado o seu parâmetro b que foi de -1,09. As medidas de correlação item-total ($r_{item-total}= 0,48$), de *Infit* (0,88), de *Outfit* (0,82) e, ainda, a medida de *Residual Loading* (0,31) também foram indicadas como adequadas ao que se propôs para esta etapa do estudo.

<p>Na aula de Biologia, em uma escola de Ensino Médio, ao trabalhar um determinado assunto a partir do livro didático adotado, o professor é interpelado por um aluno sobre a atualidade daquela matéria. O aluno explicou que, tendo acessado o site de uma universidade pela Internet, leu que havia novos conhecimentos sobre o conteúdo em pauta, que contrariavam o que estava no livro. Diante da situação, o professor, que sempre tivera posturas que valorizam a produção de conhecimentos pelos alunos, deve</p> <p>➔ (A) incentivar a turma a pesquisar sobre o assunto para avaliar as novas informações trazidas pelo aluno, deslocando a discussão para uma próxima aula.</p> <p>(B) desqualificar a Internet como meio de transmissão do conhecimento, informando sobre a existência de muitos sites não confiáveis.</p> <p>(C) acatar a informação do aluno como verdadeira, indicando à turma que esse conhecimento será objeto de avaliação.</p> <p>(D) recomendar à turma que estude pelo livro didático adotado, explicando que a prova terá o livro como base.</p> <p>(E) impor-se ao aluno, confirmando que o livro adotado é atual e suas informações estão corretas.</p>	
Conteúdo	Habilidade
Tecnologias da Comunicação	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Compreender o desenvolvimento e a aprendizagem de crianças, jovens e adultos, considerando as dimensões cognitivas, afetivas, socioculturais, éticas e estéticas.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 52.

Figura 13. Questão Correspondente ao item 17 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

A questão mencionada na Figura 13 foi considerada a mais fácil de todo o Componente Específico do ENADE de Pedagogia aplicado no ano de 2005, de acordo com os resultados divulgados pelo o Inep (2005) em que atribui-se para essa questão um índice de acerto maior que 86% e, também, de acordo com o índice de dificuldade indicado pela TRI que foi de -2,91. As demais medidas foram consonantes com os propósitos desta etapa do estudo em que, a correlação item-total angariou o valor de 0,30, o *Infit* e o *Outfit* os valores 0,94 e 0,80, respectivamente e o *Residual Loading* que obteve o valor de 0,05.

<p>Na figura abaixo, de Francesco Tonucci, a professora reage ao “erro” do aluno de forma contundente, desvalorizando o seu raciocínio analógico. O erro e sua correção tiveram, ao longo do tempo, diferentes abordagens relacionadas a concepções e reflexões sobre a avaliação da aprendizagem.</p>	
 <p>(TONUCCI, F. Caminhos de criança. Porto Alegre: Artes Médicas, 1997, p. 83)</p>	<p>O procedimento docente que caracteriza uma concepção mediadora de avaliação é:</p> <p>(A) determinar a correspondência entre as intenções e as observações, verificando se os dados observados correspondem às intenções formuladas.</p> <p>(B) respeitar as diferenças individuais e, sem comparar um aluno em relação a outros, fazer um julgamento com base nos objetivos alcançados por ele.</p> <p>→ (C) analisar as várias manifestações dos alunos em situações de aprendizagem, considerando suas hipóteses, para exercer uma ação educativa.</p> <p>(D) valorizar a produção individual sobre a coletiva, acompanhando os alunos em diversas situações de aprendizagem, para estabelecer as estratégias de ensino.</p> <p>(E) considerar o resultado obtido pelo aluno como consequência do seu empenho em reproduzir os conteúdos dados e as normas estabelecidas pela instituição onde ele estuda.</p>
<p>Conteúdo</p>	<p>Habilidades</p>
<p>Avaliação do ensino e da aprendizagem.</p>	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Compreender o desenvolvimento e a aprendizagem de crianças, jovens e adultos, considerando as dimensões cognitivas, afetivas, socioculturais, éticas e estéticas.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 52.

Figura 14. Questão Correspondente ao item 18 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

Continuando a descrição das análises e seus respectivos resultados, para essa questão (Figura 14) o valor correspondente à medida de correlação item-total foi de 0,35; já para as medidas de *Infit* e *Outfit* os valores (0,98 e 1,0, respectivamente) não evidenciaram nenhum padrão de resposta inesperado. A medida *Residual Loading* também apresentou adequação ao critério estabelecido e apresentou o valor de 0,18. Este item, segundo o parâmetro que indica o índice da dificuldade, apresentou nível médio de dificuldade, conforme $b = 0,40$.

<p>Uma estudante do Curso de Pedagogia observa, durante o estágio, uma criança utilizando um cabo de vassoura para puxar seu brinquedo que rolou para debaixo do armário e reflete sobre os diferentes aspectos do desenvolvimento, destacando o que corresponde àquela situação, a fim de registrá-la no relatório sobre a criança. Embora vários aspectos do desenvolvimento estejam presentes na situação descrita, qual é o predominante?</p> <p>(A) Social. (B) Cultural. <input checked="" type="checkbox"/> (C) Intelectual. (D) Físico-motor. (E) Afetivo-emocional.</p>	
Conteúdo	Habilidades
Psicologia da Educação	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Compreender o desenvolvimento e a aprendizagem de crianças, jovens e adultos, considerando as dimensões cognitivas, afetivas, socioculturais, éticas e estéticas.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 52.

Figura 15. Questão Correspondente ao item 20 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

Do mesmo modo, a questão descrita na Figura 15, revela um nível médio de dificuldade, haja vista o parâmetro referente à esta medida que foi de $b= 0,69$. Os demais critérios também foram privilegiados; assim, a medida da correlação item-total foi de 0,30, a de *Infit* atingiu o valor de 1,02, a de *Outfit* apresentou o valor de 1,03 e, por último, o valor correspondente à medida *Residual Loading* foi de -0,12.

<p>Duas crianças muito competitivas recusaram a possibilidade de desenvolver, juntas, uma mesma atividade. A professora incentivou os alunos a trabalhar em dupla, “para que pudessem aprender mais”. Ao fazer esta consideração, a professora se aproxima do princípio de que a aprendizagem é um processo</p> <p>(A) inato. <input checked="" type="checkbox"/> (B) social. (C) empírico. (D) “gestáltico”. (E) condicionante.</p>	
Conteúdo	Habilidades
Psicologia da Educação	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Compreender o desenvolvimento e a aprendizagem de crianças, jovens e adultos, considerando as dimensões cognitivas, afetivas, socioculturais, éticas e estéticas.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 52.

Figura 16. Questão Correspondente ao item 21 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

O item de número 21 do Componente Específico do ENADE (Figura 16) pôde ser considerado, de acordo com os resultados obtidos por meio da TRI, como o segundo item mais fácil da prova ($b= -1,18$) dentre os itens escolhidos para comporem a prova PAC. Ressalta-se que, também para este item, as medidas de correlação item-total

($r_{item-total} = 0,33$), de *Infit* (1,00), de *Outfit* (0,97) e a correspondente ao *Residual Loading* (-0,04) foram adequadas e suficientes.

<p>Ser professor significa desenvolver atividades pedagógicas e projetos político-pedagógicos, questionar a própria prática e refletir sobre o fazer profissional. Na proposta de formação do professor-pesquisador entende-se que o docente deve</p> <p>(A) vivenciar a atividade de pesquisa em disciplinas teóricas.</p> <p>(B) modificar regularmente seus pressupostos teóricos.</p> <p>→(C) organizar a sua ação a partir da articulação prática-teoria-prática.</p> <p>(D) trabalhar a teoria como um suporte complementar cuja finalidade é a prática.</p> <p>(E) condicionar sua prática no sentido de observar e alterar seu fazer pedagógico.</p>	
Conteúdo	Habilidades
<p>Constituição do “ser” professor e as especificidades do trabalho docente</p>	<p>Planejar, desenvolver e avaliar situações de ensino e de aprendizagem, de modo a adequar objetivos, conteúdos e metodologias específicos das diferentes áreas à diversidade dos alunos e aos fins da educação;</p> <p>Proceder à avaliação e organização de conteúdos e de estratégias para a aprendizagem, considerando as múltiplas dimensões da formação humana.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 53.

Figura 17. Questão Correspondente ao item 23 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

Este item (Figura 17), de acordo com a calibragem dos itens, foi considerado de média dificuldade indicado pelo valor de seu parâmetro $b = 0,23$. A correlação item-total apresentou o valor de 0,33 – adequada conforme os critérios postos; a medida de *Infit* atingiu o valor de 1,01 e, a de *Outfit* o valor de 1,02 e, a medida *Residual Loading* obteve o valor de 0,16. Assim, todas as medidas contemplaram os critérios estabelecidos.

Analisar as duas imagens a seguir, de diferentes autores e em diferentes épocas.



Roberto Doisneau.



Ilustração de Mariana Massarani,
in Ruth Rocha, Nicolau tinha uma idéia

O elemento imagético que representa a escola em ambas é a

- (A) postura irregular dos alunos.
- (B) organização espacial simétrica.
- (C) maneira de o professor se posicionar.
- (D) atenção dos alunos à fala do professor.
- (E) presença de materiais didáticos diversos.

Conteúdo		Habilidades
Linguagens Culturais.	Artístico-	Participar da formulação, implementação e avaliação contínua de projetos pedagógicos escolares e não escolares; Analisar situações educativas e de ensino e realizar estudos e pesquisas, de modo a produzir conhecimentos teóricos e práticos.

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 54.

Figura 18. Questão Correspondente ao item 28 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

O penúltimo item (Figura 18) escolhido para compor o PAC de Pedagogia foi o de número 28 do Componente Específico do ENADE. Este item também pôde ser considerado com nível médio de dificuldade dado o valor apresentado por seu parâmetro b que foi de 0,25. Também foram encontrados valores que podem ser avaliados como suficientes quais sejam, correlação item-total ($r_{item-total}= 0,39$), *Infit* (0,96), *Outfit* (0,95) e *Residual Loading* (0,18).

<p>A questão da cidadania, categoria histórica e de importância fundamental como princípio educativo, alargou seus horizontes de atuação nos anos 80, ao incorporar a problemática dos direitos coletivos e não apenas individuais. (GOHN, Maria da Glória. Movimentos sociais e educação. São Paulo: Cortez, 2001, p. 112).</p> <p>Essa afirmação pressupõe que a educação tem como um dos seus objetivos:</p> <p>→ (A) alterar modos de agir, concepções e representações na sociedade civil e política. (B) privilegiar ações corporativas segundo o interesse de grupos envolvidos. (C) propiciar movimentos educacionais para a manutenção da ordem vigente. (D) construir e aprimorar canais de representação para garantia do direito ao voto. (E) implantar linhas e diretrizes no interior dos movimentos sociais.</p>	
Conteúdo	Habilidades
Relação escola-comunidade e movimentos sociais.	<p>Conhecer a realidade dos diferentes espaços de atuação e suas relações com a sociedade, de modo a propor intervenções educativas fundamentadas em conhecimentos filosóficos, sociais, históricos, econômicos, políticos, artísticos e culturais;</p> <p>Conhecer e analisar as políticas educacionais e seus processos de implementação;</p> <p>Promover e planejar ações visando a gestão democrática nos espaços e sistemas escolares e não escolares.</p>

Fontes: Prova do Componente específico do ENADE de Pedagogia/2005; Relatório Síntese da área de Pedagogia, 2005, p. 55.

Figura 19. Questão Correspondente ao item 34 do componente de CE do ENADE de Pedagogia/2005 e habilidades relacionadas.

Por fim, o item de número 34 no Componente Específico do ENADE de Pedagogia (Figura 19), como todos os itens escolhidos para comporem o PAC de Pedagogia, excetuando-se o item de número 15 que não se adequou apenas em um critério, contemplou todos os requisitos propostos para esta etapa do estudo. Dessa maneira, a medida de correlação item-total foi de 0,38, as medidas de *Infit* e *Outfit* foram, respectivamente, de 1,03 e 1,04 e a de *Residual Loading* foi de -0,24. O nível de dificuldade atribuído para esta questão foi médio, indicado por $b = -0,18$.

4.2. Estudo 2 – Análise de uma prova PAC de Pedagogia equalizada ao ENADE 2005

4.2.1 Método

Fonte de dados

Este estudo contou com um banco de dados contendo informações referentes às respostas dos estudantes do curso de Pedagogia a uma Prova PAC 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

A amostra deste estudo foi composta por 260 estudantes matriculados no curso de Pedagogia de uma universidade particular do interior do Estado de São Paulo, sendo 93,1% mulheres, em diferentes momentos do curso, distribuídos em três grupos a saber, ingressantes, intermediários e concluintes (Tabela 17). O critério para a distribuição dos estudantes, descrita na Tabela 16, 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.

Tabela 16

Critérios para a diferenciação da situação acadêmica

Semestre do curso	Situação acadêmica	
	Carga horária cursada (%)	Categoria
1	12,5	Ingressante
2	25,0	Ingressante
3	37,5	Intermediário
4	50,0	Intermediário
5	62,5	Intermediário
6	75,0	Intermediário
7	87,5	Concluente
8	100,0	Concluente

Dessa forma, estudantes que cursaram até 25% do curso serão considerados ingressantes; àqueles que cumpriram até 75% da graduação serão observados como intermediários e, os últimos, acima de 75% do curso realizado, considerados como concluintes. A Tabela 17 mostra a distribuição de estudantes por diferenciação acadêmica, em que a maioria de 59,6% é considerada intermediária, de acordo com os critérios expostos na Tabela 16.

Tabela 17

Distribuição dos estudantes pela situação acadêmica

Situação acadêmica	Número de estudantes	%
Ingressantes	42	16,2
Intermediários	155	59,6
Concluintes	63	24,2
Total	260	100,0

Material

Seruiu de material uma prova do PAC 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, sendo 05 desses itens comuns ao ENADE de Pedagogia de 2005; 10 itens de Componente Específico referentes ao Núcleo de Conteúdos Básicos ou Fundamentais; 20 itens de Componente Específico referente ao Núcleo de Conteúdos Profissionalizantes – desses itens, 11 deles são comuns ao ENADE de Pedagogia de 2005.

Procedimento

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 as necessidades dos softwares utilizados na pesquisa.

4.2.2 Resultados e Discussão

O estudo 2 teve por objetivo a análise da prova PAC aplicada aos estudantes de Pedagogia, de uma universidade do interior de São Paulo, de acordo com os parâmetros obtidos pela TRI. Ressalta-se que, para cumprir os propósitos deste estudo, utilizou-se a mesma metodologia semelhante à etapa 1 do estudo 1, em que os dados foram tratados para que fosse possível sua utilização ao *software* estatístico *Winsteps*. Assim, a Tabela 18 fornece os resultados alcançados com a aplicação da TRI, em que, os itens âncoras são indicados pela letra “A”, logo após o índice de dificuldade do item.

Tabela 18

Estatísticas dos itens em Formação Geral no PAC de Pedagogia calibradas pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	<i>Displace</i>
01	171	209	-0,45A	0,78	0,68	0,47	-0,26
02	137	208	-0,07A	1,11	1,09	0,46	0,38
03	190	208	-1,62A	0,96	0,81	0,32	-0,08
04	107	209	1,17A	1,06	1,06	0,37	-0,12
05	163	209	-0,33A	1,04	1,06	0,33	-0,09
06	140	208	0,23	0,94	0,92	0,46	0,00
07	082	206	1,59	1,07	1,13	0,34	0,00
08	105	208	1,08	1,10	1,16	0,32	0,00
09	099	210	1,25	0,91	0,89	0,49	0,00
10	074	208	1,82	0,99	0,94	0,41	0,00
Média	126,8	208,3	0,47	1,00	0,97		
DP	37,4	1,0	1,04	0,10	0,14		

Conforme a Tabela 18, os itens escolhidos para comporem a prova PAC apresentaram, de acordo com a TRI, medidas consideradas suficientes no que concerne à qualidade dos itens – nenhuma medida ultrapassou ou foi inferior aos valores estabelecidos para a suficiência dos itens. Considerando a prova como um todo, a questão que apresentou o menor índice de dificuldade, ou seja, exigiu um nível menor de habilidade do respondente refere-se à questão 3 (âncora; $b = -1,62$). Do mesmo modo, a questão que exigiu maior habilidade por parte dos participantes foi a questão de número 10 ($b = 1,82$), retirada do componente de Formação Geral do ENADE 2008, segundo os propósitos do colegiado do curso de Pedagogia.

A Figura 20 corresponde à Curva Característica do Item 3; nessa curva é expressa no eixo X o nível de habilidade dos sujeitos e no eixo Y a probabilidade de acerto ao item,

sendo que a curva expressa a probabilidade de acerto dos participantes com os vários níveis de habilidade tendo em vista a dificuldade daquele item. Nessa figura é apresentada a resposta correta marcada na legenda ao lado direito por dois asteriscos. Por essas informações é possível visualizar quais alternativas as pessoas com diferentes níveis de habilidade marcam em itens difíceis, médios e fáceis.

Analisando a Figura 20 percebe-se que pessoas com níveis de habilidade de aproximadamente $-0,70$ respondem também as demais categorias, ou seja, respondem também às outras opções de respostas. Conforme a habilidade dos participantes aumenta, a resposta correta (A) passa a ser a mais provável para um nível de habilidade de até aproximadamente $1,2$ em que a probabilidade é $1,0$ de acerto.

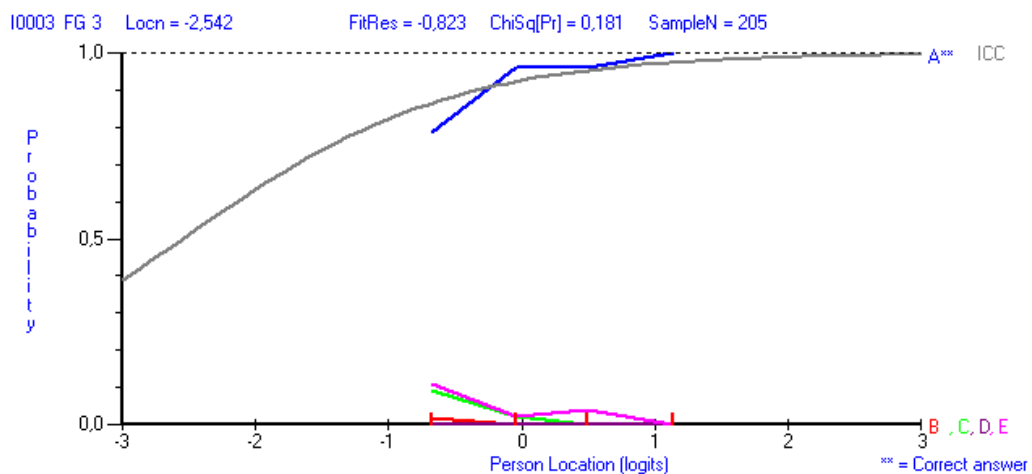


Figura 20. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 03 da prova de Formação Geral do PAC.

Já para a questão de número 10, a Figura 21 apresenta a mesma relação para o item mais difícil da prova de Formação Geral do PAC. Pode-se observar nesse caso que a

resposta correta (D) é pouco provável em praticamente todos os níveis de habilidade das pessoas. Nesse caso, essa probabilidade começa a aumentar próximo ao ponto 0 de habilidade e com pessoas mais habilidosas essa resposta passaria a ser a mais provável e poderia chegar até a probabilidade de acerto de, aproximadamente, 0,5. As categorias de resposta B e E são as mais prováveis em pessoas com menores habilidades. Assim, conclui-se que este é um item que apresenta problemas na sua construção ou que a habilidade requerida não foi a mesma envolvida nas outras questões da prova.

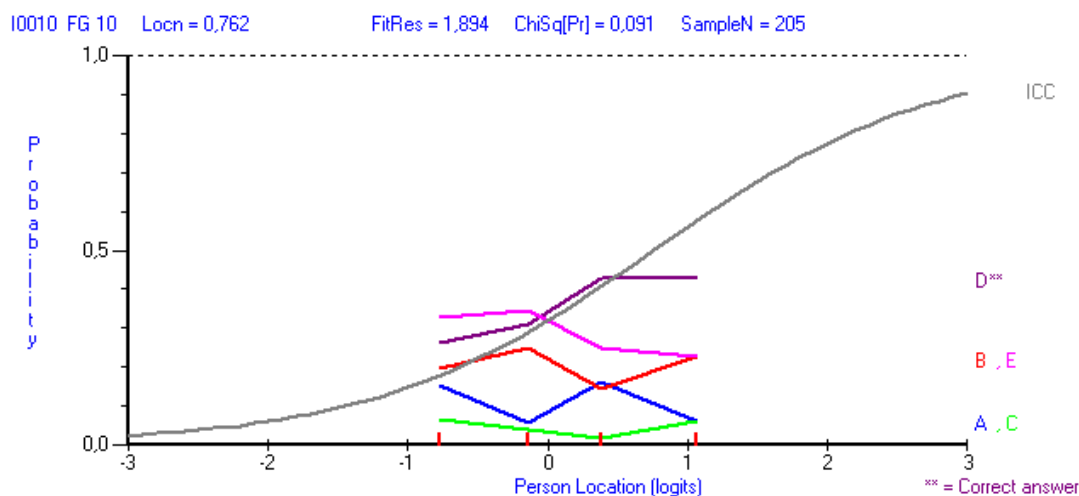


Figura 21. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 10 da prova de Formação Geral do PAC.

Com relação aos itens âncoras, indicado na Tabela 19 pela letra “A”, estes apresentaram medidas adequadas em que, a correlação item-total variou entre 0,32 e 0,47, a média *Infit* foi de 1,00 ($DP= 0,10$) e a de *Outifit* de 0,97 ($DP= 0,14$). Os itens âncoras da Formação Geral do PAC, quanto ao índice de dificuldade, se subdividiram em itens fáceis, médios e difíceis. Dentre eles, o item mais fácil apresentou índice de dificuldade de -1,62

(item 3) e, por outro lado, o item que pôde ser considerado o mais difícil apresentou índice de dificuldade de 1,17 (item 4).

Os índices de ajuste ao modelo de *Rasch* referentes aos itens do Componente Específicos do PAC Pedagogia estão dispostos na Tabela 19. Considerando a prova como um todo, isto é, não distinguindo entre itens âncoras e itens comuns, houve itens que podem ser considerados inadequados no que se refere às suas medidas. A medida de correlação item-total, por exemplo, obteve valores abaixo do mínimo esperado que é de 0,20 nas questões 27 ($r_{item-total}= 0,14$), 36 ($r_{item-total}= 0,08$) e 39 ($r_{item-total}= -0,08$). Na medida de *Outfit* também se manifestaram valores que indicam padrões de respostas inesperados como as questões 36 ($Outfit=1,44$) e 39 ($Outfit=2,10$), acima do valor crítico de 1,40.

Cabe mencionar que dois itens, anteriormente selecionados para comporem a prova de Conhecimento Específico do PAC, foram retirados das análises por apresentarem medidas de *Displace* superiores as consideradas adequadas (acima de 1 ou -1). Essa medida é uma estimativa que mede a conformidade do item com relação ao seu item âncora. Altos valores nessa medida indicam que o escore de um item observado é menor do que o escore de um item esperado, baseado na medida ancorada (Linacre & Wright, 1991).

Tabela 19

Estatísticas dos itens no Componente Específico do PAC de Pedagogia calibradas pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	<i>Displace</i>
11	064	211	0,70A	0,85	0,79	0,48	0,09
12	130	208	-0,73	1,08	1,10	0,26	0,00
13	061	211	0,70A	1,01	0,97	0,27	0,18
14	148	211	-1,09A	0,83	0,73	0,55	-0,03
15	068	211	0,69A	1,04	1,07	0,28	0,01
16	155	211	-1,18A	0,98	1,02	0,30	-0,14
17	092	210	0,25A	1,04	1,03	0,34	-0,11
18	094	211	-0,18A	1,06	1,07	0,29	0,28
19	157	210	-1,38	0,89	0,78	0,47	0,00
20	147	212	-1,09	1,02	1,07	0,30	0,00
21	057	211	0,99	1,01	1,07	0,30	0,00
22	089	212	0,22	0,82	0,80	0,57	0,00
23	110	210	-0,25	0,94	0,92	0,43	0,00
24	091	210	0,16	0,92	0,91	0,45	0,00
25	099	210	-0,28A	0,92	0,91	0,47	0,27
26	105	211	-0,14	1,02	1,01	0,35	0,00
27	042	208	1,41	1,10	1,29	0,14	0,00
28	091	209	0,16	1,09	1,17	0,24	0,00
29	143	212	-0,99	0,88	0,85	0,48	0,00
30	098	211	0,02	1,11	1,13	0,23	0,00
31	109	211	-0,22	1,09	1,08	0,26	0,00
32	181	211	-2,17	0,91	0,69	0,41	0,00
33	104	209	0,40A	1,15	1,18	0,32	-0,51
34	104	211	-0,12	1,07	1,11	0,28	0,00
35	073	209	0,57	1,06	1,06	0,28	0,00
36	041	207	1,44	1,13	1,44	0,08	0,00
37	142	210	-0,98	0,82	0,76	0,56	0,00
38	148	208	-1,17	0,97	0,93	0,38	0,00
39	021	209	2,30	1,15	2,10	-0,08	0,00
40	090	209	0,18	0,98	0,97	0,38	0,00
Média	101,8	210,1	-0,06	1,00	1,03		
DP	38,4	1,3	0,95	0,10	0,26		

A prova do Componente Específico do PAC de Pedagogia apresentou itens que, de acordo com seus respectivos índices de dificuldade, podem ser considerados muito fáceis, fáceis, médios, difíceis e muito difíceis. Na Figura 22 é possível visualizar a distribuição dos participantes da prova e os índices das dificuldades de cada item. Desse modo, pode-se dizer que as questões que exigiram um nível maior de habilidade dos respondentes foram as de número 27 ($b= 1,41$) e 36 ($b= 1,44$); Já a questão mais fácil da prova foi a de número 32 ($b= -2,17$).

Outra característica que pode ser observada na Figura 22, menciona a letra “M” de ambos os lados da figura e que indica os valores médios de habilidade das pessoas/dificuldade dos itens. Isso posto, é possível notar que o teste foi de dificuldade mediana para os sujeitos avaliados – as pessoas apresentaram uma média de habilidade próxima (um pouco abaixo) à média de dificuldade dos itens.

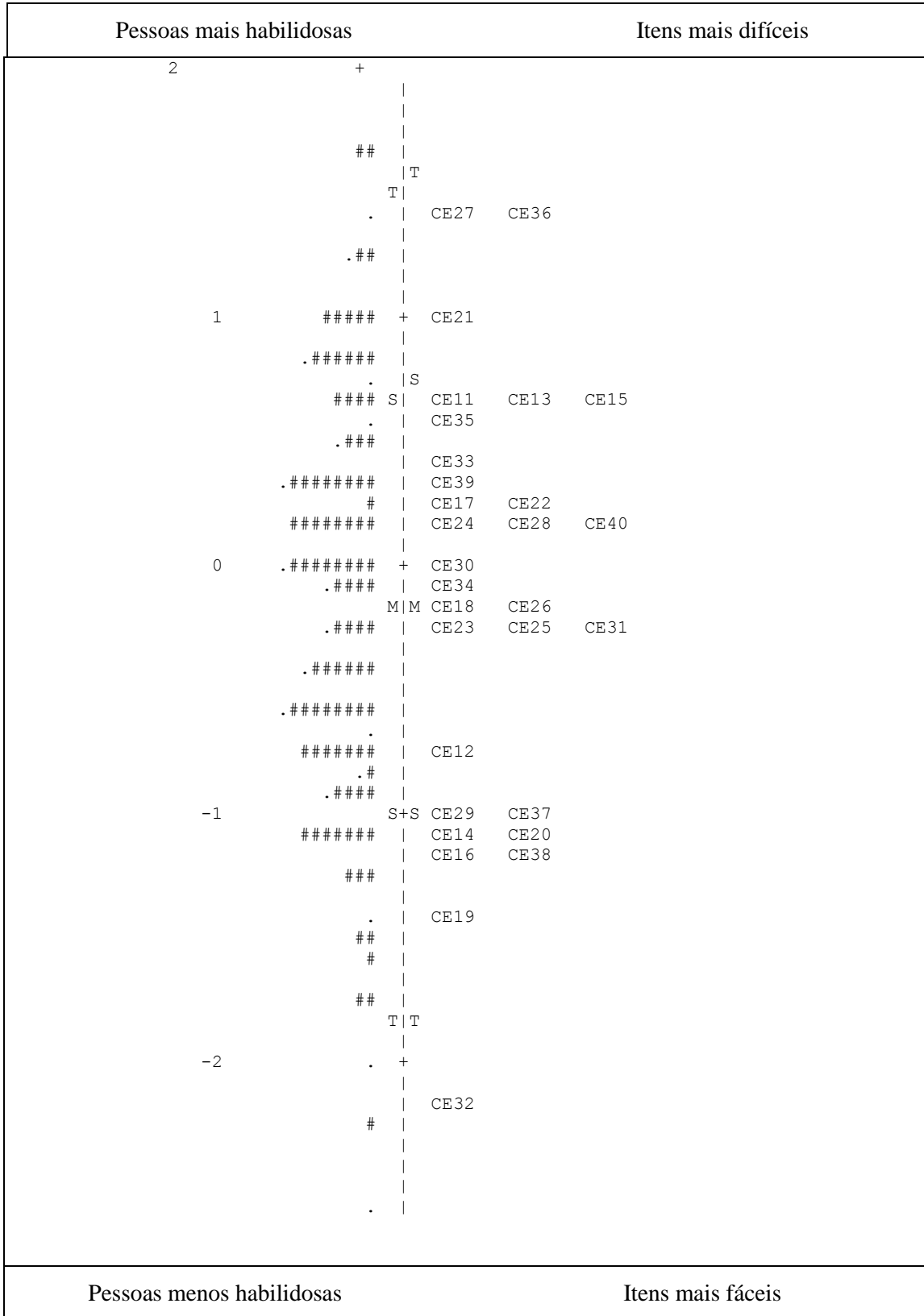


Figura 22. Comparação de habilidades dos estudantes e dificuldade dos itens.

A curva característica referente ao item mais difícil da prova é apresentada na Figura 23, em que a probabilidade de acerto ao item corresponde ao eixo Y e a habilidades dos sujeitos corresponde ao eixo X. Pode-se observar, nesse caso, que a resposta correta é pouco provável em praticamente todos os níveis de habilidade das pessoas (B). Nesse caso, essa probabilidade começa a aumentar próximo ao nível de habilidade 0. As repostas mais prováveis a partir de aproximadamente 0,5 de habilidade são dadas pela categoria E que cresce com o aumento da habilidade dos sujeitos, e a categoria B que depois de 0,5 começa a decrescer. Este é um item que requer especial atenção, uma vez que pode sugerir problemas na construção do mesmo.

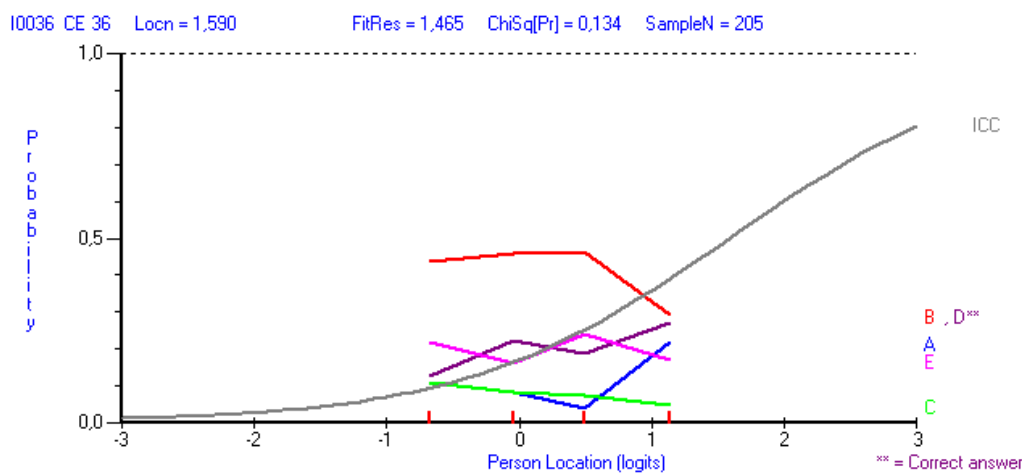


Figura 23. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 36 da prova de Componente Específico do PAC.

Para o item mais fácil da prova do Componente Específico da prova PAC, obteve-se a curva característica do item evidenciada na Figura 24. Assim, no item indicado como representante das competências mínimas a serem desenvolvidas pelo curso, percebe-se que

peessoas com níveis de habilidade de aproximadamente $-0,70$ ainda tem 25% de chance de responderem errado, com possibilidades de resposta nas categorias B e C. Conforme a habilidade das pessoas aumenta, a resposta correta (A) passa a ser a mais provável para um nível de habilidade de até aproximadamente $1,10$ em que a probabilidade é quase $1,00$ de acerto.

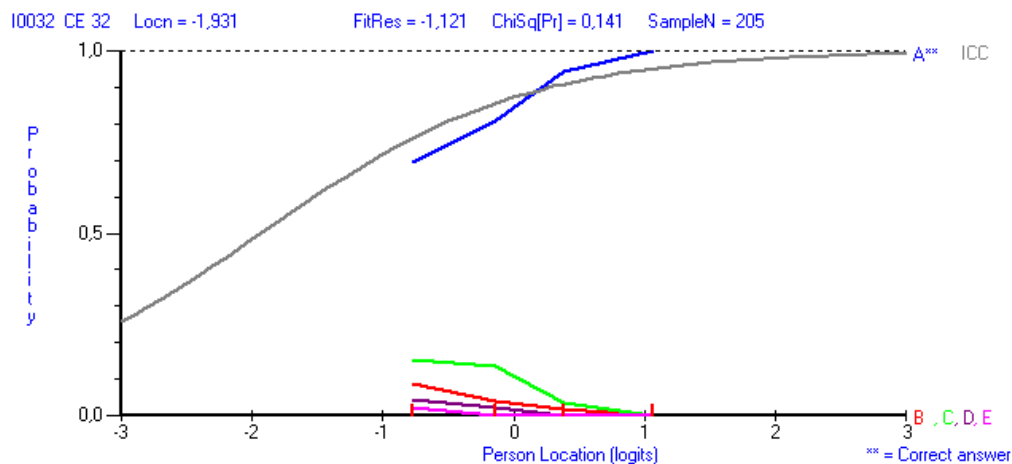


Figura 24. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 32 da prova de Componente Específico do PAC.

A Tabela 20 apresenta os parâmetros psicométricos dos itens das provas administradas no PAC de Pedagogia 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. Por esses dados, identifica-se que, em média, os itens apresentaram dificuldade mediana ($-0,13$) e houve indicadores que alcançaram até $1,44$ de dificuldade (máximo) e $-2,19$ (mínimo). Os índices de ajuste de *Infit* e *Outfit* estiveram bons, sendo que apenas um índice de *Outfit* foi igual a $1,51$, que é o

limite indicado de desajuste (1,4). Quanto às pessoas, a maior parte evidenciou habilidades semelhantes à da média dos itens (-0,16), o que indica, como já citado, que a prova possui um nível de dificuldade semelhante à habilidade das pessoas. Não houve pessoas com habilidades muito elevadas, e a mais alta alcançou o nível 1,7 de theta. Em contrapartida, houve pessoas com baixa habilidade (theta = -2,60).

Tabela 20

Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas e itens.

Parâmetro dos itens

Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	103,8	-0,13	1,00	1,00
Desvio Padrão	35,6	0,85	0,10	0,17
Máximo	181,0	1,44	1,16	1,51
Mínimo	41,0	-2,19	0,82	0,69

Parâmetros das pessoas

Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	14,7	-0,16	1,00	1,00
Desvio Padrão	5,0	0,83	0,15	0,25
Máximo	25,0	1,70	1,43	2,18
Mínimo	3,0	-2,60	0,72	0,48

Destarte os resultados do Estudo 1 e Estudo 2 acima descritos, a próxima seção pretende destacar os resultados da comparação entre o desempenho dos estudantes do ENADE de Pedagogia de 2005 e dos estudantes do PAC de Pedagogia. Sendo assim, segue o proposto.

4.3 *Estudo 3* – Análise e interpretação dos resultados do PAC comparados ao ENADE em uma mesma escala

4.3.1 Método

Fonte de dados

Para este estudo contemplou-se as Fontes de dados utilizadas nos Estudos 1 e Estudo 2 quais sejam, respectivamente, as bases do ENADE 2005 e as bases de dados do PAC de Pedagogia.

Participantes

Para esse estudo foram participantes as amostras do estudo 1, composta pelos estudantes que fizeram o ENADE de Pedagogia em 2005, e do estudo 2, composta pelos estudantes que realizaram a prova PAC de Pedagogia.

Procedimento de análise dos dados

Os dados dos dois estudos anteriores foram organizados com o objetivo de comparar as duas amostras. Para tanto, contou-se com o auxílio do *Statistical Package for Social Sciences for Windows (SPSS)*, em que as frequências de acertos dos estudantes, tanto do ENADE 2005 quanto do PAC, foram transformadas em percentis a fim de uma melhor análise dos resultados.

4.3.2 Resultados e Discussão

O presente estudo teve por objetivo apresentar os resultados das análises psicométricas do componente específico da prova aplicada ao curso de Pedagogia do PAC e compará-los aos dados do ENADE 2005, por meio de procedimentos de equalização de provas com itens comuns ao ENADE 2005 e PAC. Com isso, os dados obtidos por meio dessas análises permitiram verificar sua adequação para a classificação do nível de habilidade dos estudantes avaliados, bem como para avaliar o desempenho obtido pelos participantes deste estudo.

A fim de que sejam proporcionados os resultados desse estudo, optou-se por dividi-los em etapas, com o intuito de apresentar, de forma didática, seus conteúdos. Assim sendo, a Tabela 21 aporta as informações sobre os conceitos dos cursos do ENADE 2005, divididos entre ingressantes e concluintes, atribuídos pelo INEP, conforme descrito anteriormente. 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 21

Distribuição de estudantes de Pedagogia por conceito no ENADE 2005.

Conceito	Ingressantes		Concluintes		Total	
	N	%	N	%	N	%
1	609	02,6	706	02,7	1315	02,7
2	2785	11,9	3515	13,4	6300	12,7
3	11552	49,5	13366	51,1	24918	50,3
4	4737	20,3	5018	19,2	9755	19,7
5	301	01,3	293	01,1	594	01,2
SC	3334	14,3	3281	12,5	6615	13,4
Total	23318	100,0	26179	100,0	49497	100,0

Seguindo a descrição dos resultados, os estudantes que fizeram o PAC foram divididos por ano de curso em que estão matriculados, sendo que a categoria ingressante do ENADE corresponde aos estudantes do 1º ano do PAC; as categorias 2º e 3º anos são considerados estudantes intermediários, sem correspondência ao ENADE; e, por fim, os estudantes do 4º ano que equivale à categoria Concluintes do ENADE. Os estudantes do PAC também foram divididos por Campus em que estavam matriculados, a Tabela 22 fornece essas informações em que pôde-se observar que a maior parte dos estudantes se concentram no Campus 2 da IES e estão matriculados entre o 3º e 4º semestres, ou seja, no 2º ano do curso de Pedagogia.

Tabela 22

Distribuição de estudantes de Pedagogia que fizeram o PAC por Campus e situação acadêmica

Ano	Campus				Total	
	1		2		N	%
	N	%	N	%		
1	01	00,9	41	27,3	42	16,2
2	63	57,8	42	28,0	105	40,5
3	19	17,4	31	20,7	50	19,3
4	26	23,9	36	24,0	62	23,9
Total	109	100,0	150	100,0	259	100,0

Para a obtenção de uma escala comum entre os estudantes do ENADE de Pedagogia de 2005 e os estudantes do PAC de Pedagogia, foi empregado o procedimento de comparação normativa, em que os acertos dos estudantes são transformados em uma escala de percentis. Essa escala varia 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. 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 PAC com o 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 ENADE. Assim sendo, os percentis foram calculados a partir das notas do ENADE e, com essa equalização, os valores referentes à escala PAC correspondem ao mesmo nível de habilidade no ENADE. A distribuição dos percentis da nota média no ENADE, divididos por ingressantes e concluintes, é apresentada na Figura 25.

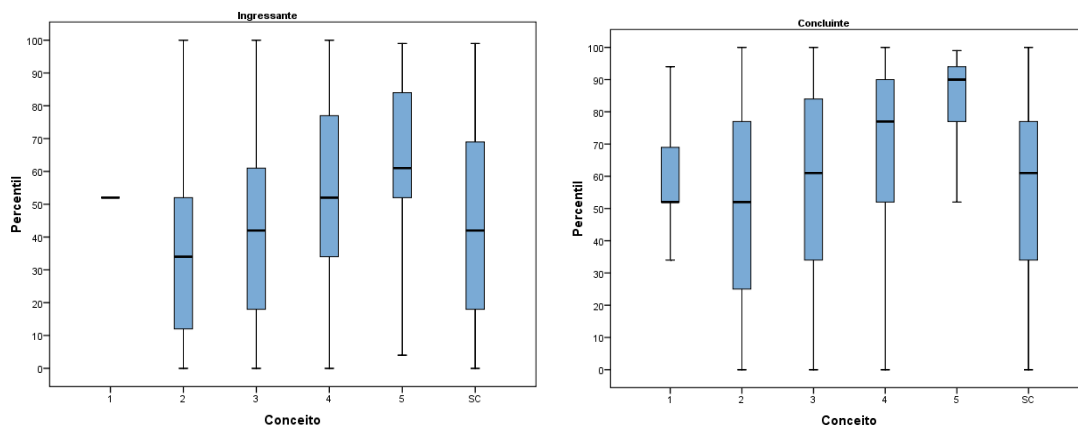


Figura 25. Box-plot dos percentis de nota média no ENADE de ingressantes e concluintes por conceito do curso.

A análise da Figura 25 permite dizer que, de acordo com os percentis obtidos, o desempenho dos estudantes que participaram do ENADE de Pedagogia 2005 aumenta conforme aumenta o conceito dos cursos. Desta forma, tantos os estudantes ingressantes quanto os estudantes 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.

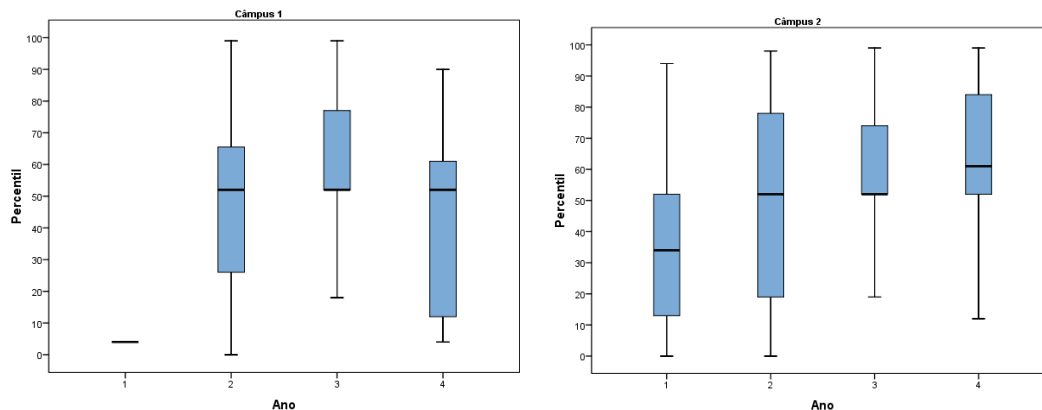


Figura 26. Box-plot dos percentis de nota média no PAC por ano e Campus.

A Figura 26 contempla os percentis de nota média dos estudantes do PAC, divididos por ano e Campus de matrícula. Conforme a figura, pode-se observar variabilidade no desempenho dos estudantes de acordo com a sequência 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 PAC 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.

Já para os estudantes do Campus 2 houve uma variabilidade consonante com os 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 desses estudantes 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 que ficou em torno do percentil 99. As notas dos dois anos de graduação foram semelhantes, mas, como pode ser visto na Figura 26, pelo menos 50% dos alunos matriculados no 4º ano obtiveram melhor desempenho no PAC.

Na comparação do desempenho entre os estudantes que fizeram o ENADE e o PAC, não puderam ser utilizadas as notas dos alunos ingressantes do PAC (matriculados no 1º ano) dos cursos do Campus 1, devido ao pequeno número de participantes. Já com relação ao desempenho dos alunos concluintes do PAC, 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 PAC, matriculados no Campus 2, quanto a equalização à escala do ENADE identificou que, de acordo com as Figuras 25 e 26, esses alunos tiveram desempenho semelhante aos estudantes ingressantes dos cursos com conceito 2 no ENADE. Considerando o desempenho dos estudantes concluintes do PAC, ou seja, matriculados no último ano de Pedagogia do Campus 2, segundo as Figuras 25 e 26, 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 Campus que fizeram o PAC de Pedagogia, observa-se que os estudantes considerados intermediários que fizeram a prova nos dois Campus apresentam desempenhos semelhantes, sendo que, 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 SAEB (Sistema Nacional de Avaliação da Educação Básica). 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 (2008).

Na avaliação do Ensino Superior que emprega o uso da equalização via TRI, destaca-se os estudos realizados pelo LabAPE (Laboratório de Avaliação Psicológica e Educacional) vinculado ao Programa de Pós-Graduação *Strico Sensu* da Universidade São Francisco. Dentre as pesquisas privilegiadas pelo LabAPE está a realização de provas institucionais construídas de modo semelhante à metodologia utilizada nesta pesquisa desde a instituição do PAC, 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 PAC, as análises psicométricas das provas PAC e, por fim, a comparação de desempenhos entre estudantes PAC e ENADE (Primi & cols., 2008a; 2008b; 2008c; 2009a; 2009b; 2009c; 2009d; 2009e; 2009f; 2009g; 2009h; 2009i; 2010a; 2010b; 2010c; 2010d; 2010e; 2010f; 2010g; 2010h; 2010i; 2010j; 2010k).

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. Silva e cols. (2011) proporcionaram um estudo cujo objetivo foi apresentar resultados de análises psicométricas de uma prova equalizada à prova de Administração do ENADE 2006, aplicada em uma IES e, assim, conferir o desempenho desses estudantes em relação aos estudantes brasileiros avaliados pelo ENADE. Foram participantes 141689 estudantes do curso de Administração que realizaram o ENADE (2006) e 916 estudantes desse mesmo curso de uma Universidade particular do interior paulista, divididos em quatro campus. O instrumento utilizado foi uma prova composta por 12 itens de Formação Geral, três deles comuns ao ENADE 2006, e por 33 itens referentes ao Conhecimento Específico, sendo 11 itens comuns ao ENADE 2006.

A análise dos dados foi feita pelo modelo de Rasch em que, por meio de um mapa de itens, pôde-se averiguar que as questões mais difíceis da prova de conhecimento específico correspondem aos itens 27 (âncora ao ENADE), 18 e 40. Na contrapartida, observou-se que o item 23 corresponde ao mais fácil da prova. Ainda de acordo com o mapa de itens, se pôde notar que os estudantes apresentaram média de habilidade inferior à média de dificuldade dos itens, indicando que a prova foi difícil para as pessoas avaliadas. Os índices de ajuste *Infit* e *Outfit* foram adequados, com médias de, aproximadamente, 1,00 (Silva & cols., 2011).

Comparando-se o desempenho dos estudantes que realizaram a prova equalizada ao ENADE, os resultados evidenciaram que o desempenho dos estudantes ingressantes dos quatro campus avaliados foi semelhante ao desempenho dos estudantes ingressantes brasileiros com Conceito 2 e 3, conceitos esses instituídos pela matriz do ENADE, cujo cálculo de obtenção já foi privilegiado nesta pesquisa. Já para os estudantes concluintes, os resultados mostraram que os estudantes do campus 2 tiveram desempenho próximo aos cursos de Conceito 4 do ENADE e, por sua vez, os estudantes do campus 3 apresentou desempenho entre os Conceitos 2 e 3, ou seja, valores médios mais baixos entre os campus da IES (Silva & cols., 2011).

Estudo semelhante foi realizado por Bartholomeu e cols. (2011) cujo objetivo foi estabelecer pontos de corte em uma prova de Direito equalizada ao ENADE 2006, em face ao procedimento de referência ao item, com o emprego do modelo de Rasch. Foram participantes dessa pesquisa 74417 estudantes do curso de Direito divididos em 73106 estudantes que realizaram o ENADE no ano de 2006 e 1311 estudantes de uma Universidade particular de ensino de dois campus distintos. Os estudantes da Universidade em questão responderam à uma prova composta por 12 itens de Formação Geral, com três itens comuns ao ENADE 2006, e por 32 itens de Componente Específico, sendo 12 deles âncoras ao ENADE 2006. As questões que não contemplaram o ENADE 2006 foram elaboradas por professores do curso de Direito da IES, baseando-se na matriz do ENADE.

Os resultados alcançados por meio da TRI evidenciaram que, em média, todas as questões apresentaram índice de dificuldade médio, com os ajustes de *Infit* e *Outfit* considerados adequados. Com relação aos estudantes, os índices indicaram que, em média, o nível de habilidade foi mediano; contudo, os ajustes de *Infit* e *Outfit* dos estudantes

tiveram valores discrepantes, em que os valores máximos foram de 1,65 e 2,49, respectivamente. Para o cálculo dos pontos de corte, isto é, valores que sugerem quais competências são necessárias na execução dos itens, os autores utilizaram um mapa de itens apresentado aos docentes da Universidade que, por sua vez, analisaram o nível de complexidade dos itens em cada faixa de dificuldade, definidos em indicativos de *alta competência*, *competência mínima* e *competência baixa* – inferior à esperada por um graduando no curso (Bartholomeu & cols., 2011).

Por meio dessas análises, os resultados puderam identificar que os itens 42, 27, 31, 36 e 45 do Componente Específico são os que sugerem *alta competência*. Os itens que aferem o nível médio de competência esperado de um aluno do curso de Direito foram os 14, 17, 33, 35, 28, 40, 41, 26, 15, 38, 39 e 13 do Componente Específico. Os demais itens foram considerados de competências mínimas esperadas desses estudantes. Prosseguindo com os resultados, teve-se que o total da amostra foi dividida em três grupos, de acordo com os pontos de corte, e submetidos à análise de seus desempenhos. Assim, pôde-se observar que cerca de 26% dos estudantes ingressantes que fizeram o ENADE 2006 e receberam Conceito 5, foram classificados no grupo considerado de *alta competência*, desempenho, esse, adequado aos estudantes concluintes. Já para os estudantes da Universidade estudada, 80% deles, dos dois campus, foram considerados na faixa de *competência baixa*. Em presença desses resultados e de discussões acerca do tema, ao autores indicam que, um dos fatores que podem ter contribuído para os resultados refere-se a falta de motivação na situação de realização da prova.

CAPÍTULO 5

CONCLUSÃO E 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. Além dessa principal proposta, este estudo também pretendeu tecer considerações sobre a questão das habilidades acadêmicas, seu percurso histórico, seus principais estudiosos e, ainda, retomar as questões acerca do modelo CHC das habilidades cognitivas, tendo em vista a premissa colocada por McGrew (2008) de que esse modelo não deve ser visto como um fim último da investigação psicométrica sobre as estruturas das capacidades humanas.

A inclusão de mais seis habilidades cognitivas possibilitou a ampliação e o aprimoramento das maneiras de se conceber a inteligência em adultos, culminando em um fértil território de pesquisas. Dentre as seis habilidades cognitivas acrescidas ao modelo CHC, a mais importante para este estudo é a *Gkn*, haja vista sua importância no desenvolvimento de competências relacionadas ao bom desempenho em tarefas acadêmicas. Assim, o sucesso em provas acadêmicas como, por exemplo, no ENADE seria altamente relacionada à domínios especializados do conhecimento, uma das características da *Gkn*.

Outra proposta para este estudo foi retomar a história da avaliação da Educação Superior no Brasil, desde as primeiras propostas e tentativas até a instituição, no ano de

2004, do Sistema Nacional de Avaliação da Educação Superior (SINAES) composto, dentre outros instrumentos de avaliação, pelo ENADE. O Ministério da Educação informa que o ENADE tem, por um de seus objetivos, avaliar as habilidades dos estudantes no que se refere ao conhecimento das áreas específicas de suas profissões e a competência na compreensão de temas específicos de suas áreas e também ligadas a outras áreas do conhecimento. Esse objetivo está presente desde a instituição do SINAES e também no Manual do ENADE do ano de 2011, disponível para consulta eletrônica (<http://www.inep.gov.br>).

Ao longo de oito anos desde sua primeira aplicação o ENADE angariou diversas críticas, como as já explicitadas neste trabalho, que indicam que os objetivos delimitados para esse instrumento, pouco a pouco, se perderam. Seja por conta de algumas Instituições de Ensino Superior (IES) utilizarem os índices do ENADE como chamativo propagandístico na captação de novos estudantes, seja pelo fato de mudanças no cenário político-econômico promoverem alterações nos sistemas maiores de avaliação. Um pequeno recorte dessa discussão também foi privilegiado nesta pesquisa.

Retomando o objetivo principal deste estudo, pôde-se averiguar que o modelo matemático proposto pela TRI mostrou-se bem sucedido para a equalização de provas de habilidades acadêmicas. Os resultados do estudo 1 indicaram que a calibração dos itens por meio do modelo de Rasch, ou de um parâmetro, foram satisfatórias tanto para a identificação das habilidades dos sujeitos e das dificuldades dos itens, quanto para a escolha de itens âncoras que compuseram a prova de Pedagogia do Programa de Avaliação Continuada (PAC).

Do mesmo modo, os resultados do estudo 2 puderam identificar a dificuldade da prova PAC de Pedagogia, os itens mais fáceis e mais difíceis e, ainda, o desempenho dos estudantes que realizaram essa prova. Por fim, os resultados do estudo 3, isto é, o produto da equalização propriamente dita, permitiu a comparação entre os estudantes que fizeram a prova PAC de Pedagogia e àqueles que fizeram a prova de Pedagogia do ENADE no ano de 2005. Esses resultados se fazem pertinentes em relação aos objetivos estipulados para o PAC 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, com esse contexto ainda em construção.

REFERÊNCIAS

- Ackerman, P. L., & Rolhus, E. L. (1999). The locus of adult intelligence: knowledge, abilities, and nonability traits. *Psychology and Aging, 14*(2), 314-330.
- Ackerman, P. L., & Lohman, D. F. (2006). Individual differences in cognitive function. In: Patricia A. Alexander & Philip H. Winne (Orgs.). *Handbook of educational psychology*, (pp. 139-160). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers
- Almeida, L. S. (1994). *Inteligência: definição e medida*. Porto: Instituto Nacional de Investigação Científica.
- Anastasi, A., & Urbina, S. (2000). *Testagem psicológica*. (7a. ed.). Porto Alegre: Artes Médicas.
- 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.
- Andriola, W. B. (1998). A avaliação da aprendizagem: uma análise descritiva segundo a Teoria de Resposta ao Item (TRI). *Educação em Debate, 20*(36), 93-102.
- 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.
- Bird, S. M., Cox, S. D., Farewell, V. T., Goldstein, H., Hlt, T., & Smith, P. C. (2005). Performance indicators: good, bad, and ugly. *Journal of Royal Statistical Society*, *168*(1), 1-27.
- Brasil (1995). Lei nº 9.131, de 24 de novembro de 1995. Altera dispositivos da Lei nº 4.024, de 20 de dezembro de 1961, e dá outras providências. *Diário Oficial da República Federativa do Brasil*. Disponível em: <http://www.planalto.gov.br/civil/LEIS/L9131.htm>. Recuperado em janeiro de 2008.
- Brito, M. R. F. (2008). O SINAES e o ENADE: da concepção à implantação. *Revista Avaliação*, *13*(3), 841-850.
- Brito, M. R., Munhoz, A., Primi, R., Gonçalves, M. H., Rezi, V., Neves, L. F., Sanches, M. H., & Marinheiro, F. B. (2000). Exames Nacionais: uma análise do ENEM aplicado à Matemática. *Revista Avaliação*, *5*(4), 445-454.
- Carroll, J. B. (1992). Cognitive abilities: The State of the Art. *Psychological Science*, *3*(5), 266-271.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- 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.

- Dias Sobrinho, J. (2010). Avaliação e transformações da Educação Superior Brasileira (1995-2009): do Provão ao SINAES. *Avaliação, 15*(1), 195-224.
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum.
- Ferrão, M. E. (2003). *Introdução aos modelos de regressão multinível em Educação*. São Paulo: Komedi.B
- Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler intelligence scales and Gf-Gc theory: A contemporary approach to interpretation*. Boston: Allyn and Bacon.
- Goldstein, H. (1998). The importance of educational research. *Improving Schools, 1*(2), 33-34.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de dados*. Porto Alegre: Bookman.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2004). *SINAES – Sistema Nacional de Avaliação da Educação Superior: Da concepção à regulamentação*. Brasília: Inep.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2005). *Relatório Síntese da Área de Pedagogia*. Brasília: Inep.
- 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.

- Limana, A., & Brito, M. R. F. (2005). O Modelo de Avaliação Dinâmica e o Desenvolvimento de Competências: algumas considerações a respeito do ENADE. *Avaliação-Revista da Rede de Avaliação Institucional da Educação Superior*, 10(2), 9-32.
- Linacre, J. M., & Wright, B. D. (1991). *WINSTEPS – Rasch-Model computer programs*. Chicago: MESA Press.
- Locatelli, I. (2002). O Sistema de Avaliação da Educação brasileira. Em: La experiência brasileña em la Educacion Básica. *IV Simpósio Iberoamericano de Investigación y Educación, I Simposio Iberoamericano de Investigación sobre Eficacia Escolar y mejora de la escuela*, Habana, Cuba.
- Lopes, F. L. (2009). *Formação Geral de estudantes de Pedagogia avaliada pelo ENADE 2005*. Relatório de Iniciação Científica, Programa de Pós-Graduação *Strictu Sensu* em Psicologia da Universidade São Francisco, Itatiba, SP.
- Lord, F. M. (1980). *Applications of Item Response Theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates Inc., Publishers.
- McGrew, K. S. (2004). *Cattell-Horn-Carroll (CHC) definition Project*. Disponível em: <http://www.iapsych.com>. Recuperado em 23/08/2010.
- McGrew, K. S., & Evans, J. J. (2004). *Internal and external factorial extensions to the Cattell-Horn-Carroll (CHC) – Theory of Cognitive Abilities: A review of fator analytic research*. Disponível em: <http://www.iapsych.com>. Recuperado em 23/08/2010.

- McGrew, K. S. (2008). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 10*(1016), 1-10.
- Ministério da Educação e Cultura. *Relatório Técnico*. Instituto Nacional de Estudos e pesquisas Educacionais Anísio Teixeira (INEP), 2005. Disponível em: <<http://www.inep.gov.br/download/enade/2005/relatorios/>>. Recuperado em abril de 2009.
- Muñiz, J. (2010). Las teorías de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo, 31*(1), 57-66.
- Muñiz, J., & Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación, 5*, 13-25.
- Nogueira, S. O. (2008). *ENADE: Análise de itens de Formação Geral e de Estatística pela TRI*. Dissertação de Mestrado, Programa de Pós-Graduação *Strictu Sensu* em Psicologia da Universidade São Francisco, Itatiba, SP.
- Nunes, C. H. S. S., & Primi, R. (2005). Impacto do tamanho da amostra na calibração de itens e estimativa de escores por Teoria de Resposta ao Item. *Avaliação Psicológica, 4*(2), 141-153.
- Nunes, C. H. S. S., & Primi, R. (2009). Teoria de Resposta ao Item: Conceitos e aplicações na Psicologia e na Educação. Em: C. S. Hutz (Org.). *Avanços e polêmicas em avaliação psicológica*, (pp. 25-69). São Paulo: Casa do Psicólogo.

- Oliveira, K. S. (2006). *Avaliação do Exame Nacional de Desempenho do Estudante pela Teoria de Resposta ao Item*. Dissertação de Mestrado, Programa de Pós-Graduação *Strictu Sensu* em Psicologia da Universidade São Francisco, Itatiba, SP.
- Paiva, G. S. (2008). Avaliação do desempenho dos estudantes da educação superior: a questão da equidade e obrigatoriedade no Provão e ENADE. *Ensaio: avaliação das políticas públicas educacionais*, 16(58), 31-46.
- Paro, D. L. F. P. (2006). *Concepções implícitas de inteligência em uma amostra de estudantes universitários do estado de São Paulo*. Tese de Doutorado, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, SP.
- Pasquali, L. (2005). Provão (ENC) de Psicologia 2000 e 2001: análise dos parâmetros psicométricos. Em R. Primi (Org.). *Temas em Avaliação Psicológica*, (pp. 289-318). São Paulo: Casa do Psicólogo.
- Pasquali, L., & Primi, R. (2007). Fundamentos da Teoria da Resposta ao Item – TRI. Em L. Pasquali (Org.). *Teoria de resposta ao Item: TRI*, (11-28). Brasília, Distrito Federal: Editora UNB.
- Peña-Suárez, E., Fernández-Alonso, R., & Fernández, J. M. (2009). Estimación del valor añadido de los centros escolares. *Aula Abierta*, 37(1), 3-18.
- Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as Additive Conjoint Measurement. *Applied Psychological Measurement*, 3(2), 237-255.

- Polidori, M. M., Marinho-Araujo, C. M., & Barreyro, G. B. (2006). SINAES: perspectivas e desafios na avaliação da educação superior brasileira. *Ensaio: avaliação das políticas públicas educacionais*, 14(53), 425-436.
- Primi, R., Santos, A. A. A., Vendramini, C. M. M., Taxa, F., Muller, F. A., Lukjanenko, M. F., & Sampaio, I. S. (2001). Competências e habilidades cognitivas: diferentes definições do mesmo construto. *Psicologia: Teoria e Pesquisa*, 17(2), 151-159.
- Primi, R. (2002). Inteligência Fluida: definição fatorial, cognitiva e neuropsicológica. *Cadernos de Psicologia e Educação - Paidéia*, 12(2/3), 57-77.
- Primi, R., Santos, A. A. A., & Vendramini, C. M. V (2002). Habilidades básicas e desempenho acadêmico em universitários ingressantes. *Estudos de Psicologia*, 7(1), 47-55.
- 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., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., Carvalho, L. F., & Miguel, F. K. (2008a). *Análise psicométrica e resultados do curso de Administração USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., Carvalho, L. F., & Miguel, F. K. (2008b). *Análise psicométrica e resultados do curso de Direito USF*. Relatório Técnico Institucional não publicado.

- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., Carvalho, L. F., & Miguel, F. K. (2008c). *Análise psicométrica e resultados do curso de Psicologia USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Carvalho, L. F., Miguel, F. K., & Vendramini, C. M. M. (2009). 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. (2009). 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., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009a). *Análise psicométrica e resultados do curso de Ciências Biológicas USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009b). *Análise psicométrica e resultados do curso de Educação Física USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009c). *Análise psicométrica e resultados do curso de Enfermagem USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009d). *Análise psicométrica e resultados do curso de Farmácia USF*. Relatório Técnico Institucional não publicado.

- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009e). *Análise psicométrica e resultados do curso de Fisioterapia USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009f). *Análise psicométrica e resultados do curso de Gestão de Turismo USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009g). *Análise psicométrica e resultados do curso de Medicina USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009h). *Análise psicométrica e resultados do curso de Nutrição USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Vendramini, C. M. M., Nunes, C. H. S. S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A. S. (2009i). *Análise psicométrica e resultados do curso de Odontologia USF*. Relatório Técnico Institucional não publicado.
- Primi, R., Carvalho, L. F., Miguel, F. K., & Silva, M. C. R. (2010). Análise do funcionamento diferencial dos itens do Exame Nacional do Estudante (ENADE) de psicologia de 2006. *Psico-USF*, 15(3), 379-393.
- 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. (2010a). *Análise Psicométrica e Resultados dos Alunos da USF Arquitetura*. Relatório Técnico Institucional não publicado.

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. (2010b). *Análise Psicométrica e Resultados dos Alunos da USF Ciência da Computação*. Relatório Técnico Institucional não publicado.

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. (2010c). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia da Computação*. Relatório Técnico Institucional não publicado.

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. (2010d). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia Elétrica*. Relatório Técnico Institucional não publicado.

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. (2010e). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia Industrial*. Relatório Técnico Institucional não publicado.

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. (2010f). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia Mecânica*. Relatório Técnico Institucional não publicado.

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. (2010g). *Análise Psicométrica e*

Resultados dos Alunos da USF Engenharia de Produção. Relatório Técnico Institucional não publicado.

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. (2010h). *Análise Psicométrica e Resultados dos Alunos da USF Química*. Relatório Técnico Institucional não publicado.

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. (2010i). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia Civil*. Relatório Técnico Institucional não publicado.

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. (2010j). *Análise Psicométrica e Resultados dos Alunos da USF Engenharia Ambiental*. Relatório Técnico Institucional não publicado.

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. (2010k). *Análise Psicométrica e Resultados dos Alunos da USF Pedagogia*. Relatório Técnico Institucional não publicado.

Raudenbush, S. W. (2005). Learning from attempts to improve schooling: the contribution of methodological diversity. *Educational Researcher*, 34(4), 25-31.

Ribeiro, I. (1998). *Mudanças no desempenho e na estrutura das aptidões: Contributos para o estudo da diferenciação cognitiva em jovens*. Braga: Universidade do Minho.

- Rolfhus, E. L., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: knowledge, intelligence, and related traits. *Journal of Educational Psychology*, 9(3), 511-526.
- 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, W., & Soares, T. M. (2010). Eficácia dos processos de linkagem na avaliação educacional em larga escala. *Estudos em Avaliação Educacional*, 21(45), 191-211.
- 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.
- Sistema Nacional de Avaliação da Educação Superior – SINAES (2006). *Resumo Técnico ENADE 2005*. DF, Brasília: Inep/MEC – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
- Schelini, P. W. (2006). Teoria das inteligências fluída e cristalizada: início e evolução. *Estudos de Psicologia*, 11(3), 323-332.
- Soares, J. F., Ribeiro, L. M., & Castro, C. M. (2001). Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de direito, administração e engenharia civil. *Dados*, 44(2), 363-396.

- 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.
- Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.
- Valle, R. C. (2000). Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, 21, 7-91.
- Vendramini, C. M. M. (2000). *Implicações das atitudes e das habilidades matemáticas na aprendizagem dos conceitos de Estatística*. Tese de Doutorado, Universidade Estadual de Campinas, Campinas, SP.
- Vendramini, C. M. M., & Noronha, A. P. P. (2003). Análise de itens na construção de testes de inteligência. *Psicologia: Teoria e Prática*, 1, 73-81.
- Vendramini, C. M. M. (2005a). Avaliação multidimensional de Desempenho do Estudante. *Revista da Rede de Avaliação Institucional da Educação Superior*, 10(3), 27-40.
- Vendramini, C. M. M. (2005b). 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.
- Vendramini, C. M. M. (2009). Análise Fatorial baseada na TRI do componente de Formação Geral do ENADE. *Revista de Educação ANEC*, 38(150), 153-163.
- Vendramini, C. M. M., Silva, M. C. R., & Dias, A. S. (2009). Avaliação de Atitudes de estudantes de Psicologia via modelo de crédito parcial da TRI. *PsicoUsf*, 14(2), 192-211.

Verhine, R. E., Dantas, L. M. V., & Soares, J. F. (2006). Do Provão ao ENADE: uma análise comparativa dos exames nacionais utilizados no Ensino superior Brasileiro. *Ensaio: avaliação de políticas públicas Educacionais*, 14(52), 291-310.