

ADRIANO LOPES DE LIMA



AVALIAÇÃO DE PROVAS DE HABILIDADES ACADÊMICAS
EM COMPUTAÇÃO: EQUALIZAÇÃO COM O ENADE VIA TRI

ORIENTADORA: PROFA. DRA. CLAUDETTE MARIA MEDEIROS VENDRAMINI

ITATIBA
2012

ADRIANO LOPES DE LIMA



AVALIAÇÃO DE PROVAS DE HABILIDADES ACADÊMICAS
EM COMPUTAÇÃO: EQUALIZAÇÃO COM O ENADE VIA TRI

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

ORIENTADORA: PROFA. DRA. CLAUDETTE MARIA MEDEIROS VENDRAMINI

ITATIBA

2012

378.1.001 Lima, Adriano Lopes de.
L696a Avaliação de provas de habilidades acadêmicas em
computação: equalização com o ENADE via TRI /
Adriano Lopes de Lima. -- Itatiba, 2012.
100 p.

Dissertação (mestrado) – Programa de Pós-Graduação
Stricto Sensu em Psicologia da Universidade São
Francisco.
Orientação de: Claudette Maria Medeiros Vendramini.

1. Avaliação em larga escala. 2. ENADE. 3. Ensino
Superior. 4. Teoria de resposta ao item. I. Vendramini,
Claudete 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

Adriano Lopes de Lima defendeu a dissertação "AVALIAÇÃO DE PROVAS DE HABILIDADES ACADÊMICAS EM COMPUTAÇÃO: EQUALIZAÇÃO COM O ENADE VIA TRI" aprovada pelo Programa de Pós-Graduação *Stricto Sensu* em Psicologia da Universidade São Francisco em 20 de dezembro de 2012 pela Banca Examinadora constituída por:

Prof. Dra. Claudette Maria Medeiros Vendramini
Orientadora e Presidente

Prof. Dra. Marjorie Cristina Rocha da Silva
Examinadora

Prof. Dra. Maria Cristina Rodrigues Azevedo Joly
Examinadora

*Dedico este trabalho às mulheres da minha vida: ...
Minha mãe, guerreira e maior incentivadora;
Minha avó, exemplo de vida e que deixou muita saudade;
Minha esposa, simplicidade e atitudes positivas;
Minha filha, razão principal do meu viver!*

AGRADECIMENTOS

Primeiramente quero agradecer a Deus, pois sem Ele nada seria possível.

À minha família, pois representa a base de tudo, origem, existência e motivos para prosseguir.

Especialmente à minha mãe, Cleonice, maior responsável por todos os meus sucessos.

À minha avó Isaura, que durante este percurso nos deixou e está ao lado de Deus nos iluminando e sempre ao meu lado.

À minha esposa Silvana, que sempre me acompanhou neste trajeto e esteve ao meu lado, inclusive nos momentos mais difíceis.

À minha filha Jéssica, que mesmo com apenas 7 anos soube me apoiar e entender as horas intermináveis em frente ao computador.

Ao Eduardo Manuel Bartalini Gallego, grande parceiro e grande amigo para todos os momentos.

À minha querida e especial orientadora, profa. Dra. Claudette, pois sem ela nada disso teria acontecido.

Ao prof. Dr. Ricardo Primi, sempre fonte de inspiração para desafios.

Às professoras Ana Paula, Acácia e Cristina Joly, pelo privilégio do convívio e por contribuírem para o meu crescimento.

À Fernanda Lopes, que muito contribui para o meu crescimento, auxiliando em minhas dúvidas.

À Universidade São Francisco, que é minha segunda casa e me acolheu há mais de 21 anos, me permitindo essa oportunidade de evolução.

A todos os funcionários da Universidade São Francisco, representados pela Ana e Paula da Secretaria de Pós-Graduação.

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

RESUMO

Lima, A. L. (2012). *Avaliação de provas de habilidades acadêmicas em Computação: equalização com o ENADE via TRI*. Dissertação de Mestrado. Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco.

Instrumentos de avaliação educacional buscam em sua essência aferir os resultados obtidos pelos indivíduos no processo de geração do conhecimento. Por meio de avaliações em larga escala, o Ministério da Educação (MEC) avalia habilidades de estudantes e qualidade de instituições de ensino, nos variados níveis de ensino. No contexto educacional, um dos instrumentos utilizado para esse fim, no nível de ensino superior, é o Exame Nacional de Desempenho do Estudante – ENADE, instituído pelo MEC no ano de 2004. O presente estudo teve como objetivo equalizar uma prova via Teoria de Resposta ao Item – TRI a partir desse exame, para o curso de Ciência da Computação. Para que isso fosse possível, o autor utilizou um banco de dados contendo informações acadêmicas de 13.955 estudantes do curso de Ciência da Computação que realizaram o ENADE em 2005 e um banco de dados contendo informações acadêmicas de 60 estudantes de uma Instituição de Ensino Superior – IES particular do interior do estado de São Paulo que realizaram uma prova de Ciência da Computação de um Programa de Avaliação Continuada – PAC. Foram realizados três estudos: o primeiro contemplou a análise e seleção dos itens da prova de Computação do ENADE 2005, o segundo apresentou a análise da prova PAC equalizada com o ENADE e o terceiro foi referente à análise e interpretação do mapa de itens e de sujeitos. O primeiro estudo permitiu observar que alguns itens apresentaram boas propriedades psicométricas, tendo sido selecionados oito deles para servir como âncoras da prova PAC. No segundo estudo foi possível observar que os itens foram, em média, de dificuldade mediana (-0,11). Com relação às pessoas, a maioria demonstrou um nível de habilidade um pouco abaixo do nível de dificuldade dos itens (-0,88), sugerindo que a prova estava com um nível de dificuldade superior ao de habilidade dos sujeitos. Ainda se pode concluir que o conceito obtido pelos alunos do PAC alcançaram notas semelhantes aos alunos do ENADE 2005, com conceito 3. Para o terceiro estudo, a análise do mapa do item sugeriu que alguns itens apresentaram problemas na sua construção e outros demonstraram que possuíam um nível de dificuldade considerada difícil, média e fácil ao medir os diferentes níveis de habilidades dos estudantes. Estudos como o apresentado neste trabalho podem contribuir para entender melhor as habilidades desenvolvidas nos estudantes e a qualidade das instituições de ensino. Esses estudos são de grande relevância para a evolução do sistema educacional de um país.

ABSTRACT

Lima, A. L. (2012). *Evaluation tests of academic skills in Computing: equalization with ENADE via TRI*. Dissertation. Post-graduate studies in Psychology, San Francisco University.

Educational assessment tools seek to essentially assess the results obtained by individuals in the process of knowledge generation. Through large-scale evaluations, the Ministry of Education (MEC) assesses student skills and quality of educational institutions, in various levels of education. In the educational context, one of the instruments used for this purpose, at the level of higher education, is the National Student Performance Exam-ENADE, established by MEC in 2004. The present study aimed to equalize a proof via Item Response Theory-IRT from that examination, for the course of Computer Science. For this to be possible, the author used a database containing information of 13,955 academic students of computer science who conducted the ENADE in 2005 and a database containing academic information of 60 students at a private higher education Institution in the interior of the state of São Paulo, who took a Computer Science test of a Continuous Evaluation Program-PAC. Three studies were performed, the first contemplated the analysis and selection of the items of the Computing Exam of the ENADE 2005, the second study presented the analysis of the test-PAC equalized with the grades of exam-ENADE, and the third study was referring to the analysis and interpretation of a map of items and subject. The first study allowed observing that some items had good psychometric properties of which eight of them were selected to serve as anchors of the test- PAC. The second study was possible to observe that the items were on average, median difficulty (-0.11). Regarding people, most demonstrated a skill level slightly below the level of difficulty of the items (-0.88), suggesting that the test had a higher level of difficulty than the skill level of the subjects, can still be concluded that the concept obtained by students of the PAC reached similar grades to the ENADE students with concept 3. For the third study, the analyzes of the item map suggested that some items have shown problems in its construction and other items showed that had a level of difficulty considered easy, medium or difficult that demonstrated to measure different students' abilities level. The studies presented in this paper can contribute to a better understanding of the skills developed in the students and the quality of educational institutions. These studies are of great importance for the evolution of the educational system of a country.

SUMÁRIO

LISTA DE FIGURAS	xi
LISTA DE TABELAS.....	xii
APRESENTAÇÃO.....	1
<u>CAPÍTULO 1 - SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO NO BRASIL.....</u>	4
CAPÍTULO 2 - HABILIDADES E COMPETÊNCIAS AVALIADAS PELOS EXAMES NACIONAIS.....	19
CAPÍTULO 3 - TEORIA DE RESPOSTA AO ITEM NA AVALIAÇÃO EDUCACIONAL	31
CAPÍTULO 4 - UM SISTEMA INSTITUCIONAL DE AVALIAÇÃO	39
OBJETIVO GERAL	46
OBJETIVOS ESPECÍFICOS	46
CAPÍTULO 5 - RESULTADOS E DISCUSSÃO.....	47
CONCLUSÃO E CONSIDERAÇÕES FINAIS	89
REFERÊNCIAS	92

LISTA DE FIGURAS

Figura 1. Estrutura hierárquica das habilidades cognitivas de Cattell-Horn-Carroll adaptada de McGrew (2004)	21
Figura 2. Curva característica do item (CCI)	35
Figura 3. Comparação de habilidades dos estudantes e dificuldade dos itens	60
Figura 4. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 3 da prova de Formação Geral do PAC	67
Figura 5. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 8 da prova de Formação Geral do PAC	68
Figura 6. Comparação de habilidades dos estudantes e dificuldade dos itens	71
Figura 7. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 13 da prova de Componente Específico do PAC	72
Figura 8. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 31 da prova de Componente Específico do PAC	73
Figura 9. Box-plot dos percentis de nota média no ENADE de ingressantes e concluintes por conceito do curso	77
Figura 10. Box-plot dos percentis de nota média no PAC por ano	78
Figura 11. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 13 da prova de Componente Específico	81
Figura 12. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 16 da prova de Componente Específico	82
Figura 13. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 36 da prova de Componente Específico	83
Figura 14. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 40 da prova de Componente Específico	84
Figura 15. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 28 da prova de Componente Específico	85
Figura 16. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 25 da prova de Componente Específico	86
Figura 17. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 21 da prova de Componente Específico	87
Figura 18. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 31 da prova de Componente Específico	88

LISTA DE TABELAS

Tabela 1. Descrição dos dez fatores amplos do modelo CHC e habilidades relacionadas	23
Tabela 2. Diferenças entre a Teoria Clássica e a Teoria de Resposta ao Item ...	32
Tabela 3. Distribuição de estudantes no ENADE 2005 de Computação por categoria administrativa	48
Tabela 4. Estatísticas básicas das questões objetivas de Formação Geral por grupo de estudantes	52
Tabela 5. Classificação das questões objetivas de Formação Geral segundo o índice de facilidade	53
Tabela 6. Estatísticas básicas em Componente Específico por grupo de estudantes	53
Tabela 7. Classificação das questões objetivas de Conhecimento Específico segundo o índice de facilidade	54
Tabela 8. Distribuição dos conceitos	55
Tabela 9. Número e percentual de cursos participantes por grandes regiões segundo o conceito obtido	55
Tabela 10. Estatísticas dos itens em Formação Geral no ENADE de Computação 2005 calibrados pelo modelo de Rasch	56
Tabela 11. Estatísticas dos itens no Componente Específico do ENADE de Computação 2005 calibradas	59
Tabela 12. Sumário das estatísticas fornecidas pelo modelo de Rasch por itens .	61
Tabela 13. Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas	62
Tabela 14. Critérios para a diferenciação da situação acadêmica	64
Tabela 15. Distribuição dos estudantes pela situação acadêmica	64
Tabela 16. Estatísticas dos itens de Formação Geral no PAC de Ciência da Computação calibrados pelo modelo de Rasch	66
Tabela 17. Estatísticas dos itens no Componente Específico do PAC de Ciência da Computação calibrados pelo modelo de Rasch	69
Tabela 18. Sumário das estatísticas fornecidas pelo modelo de Rasch dos itens .	74
Tabela 19. Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas	74
Tabela 20. Distribuição de estudantes de Ciência da Computação por conceito no ENADE 2005	76

APRESENTAÇÃO

Com nove anos de existência, o Sistema Nacional de Avaliação da Educação Superior (SINAES) vem contribuindo, por meio de seus resultados, com as pesquisas na área da Psicologia e Educação. Os dados oriundos de suas aplicações permitem que instrumentos de avaliação sejam aprimorados por meio de estudos estatísticos e psicométricos.

O desenvolvimento de instrumentos de medidas, pautados em métodos psicométricos, representam, nas diversas áreas do saber, comprovada importância. Isso não é diferente quando focamos a questão educacional.

No Brasil, onde diversas etnias e culturas convivem sob um sistema educacional regulado pelo Ministério da Educação, e em que as políticas públicas são aplicadas de uma maneira generalista, o que não permite a diferenciação necessária entre as mesmas, o progresso da educação pode ser mais difícil de ocorrer. Assim, estudos referentes à avaliação educacional são importantes para o desenvolvimento do próprio país.

Motivado por essa premissa, o presente estudo focou o ensino superior e, para melhor compreensão, a pesquisa foi dividida em capítulos. A estrutura adotada tem início pelo cenário dos sistemas de avaliação da educação no país.

Dessa forma, esse contexto permite uma ampla discussão sobre quais variáveis melhor explicam o desenvolvimento educacional no país e que impliquem propostas de melhoria de um sistema educacional que caminha para um desenvolvimento cultural e para a projeção de um ensino de qualidade nos diversos níveis da educação. E assim, para que os resultados sejam conseguidos, o governo tem adotado avaliações em larga escala em seus

processos avaliativos. O Ministério da Educação (MEC) utiliza a avaliação em larga escala para a Educação Básica, o Ensino Médio e o Ensino Superior, já consolidada no Brasil e que vem se aprimorando nos últimos anos.

Em busca da qualidade no ensino superior, o país utilizou, com o transcorrer dos anos, diversos mecanismos e instrumentos de avaliação, que contribuíram para o desenvolvimento do modelo atual, o Sistema Nacional de Avaliação da Educação Superior (SINAES). Um de seus componentes, o Exame Nacional de Desempenho do Estudante (ENADE), que é objeto do presente estudo, tem como objetivo avaliar o estudante e, a partir de seus resultados, juntamente com outros indicadores de professores e infraestrutura institucional, atribuir um conceito para os cursos, elaborando assim indicadores de qualidade e ações pontuais aos cursos cujo desempenho fiquem abaixo do padrão esperado.

As instituições de ensino superior (IES), atentas a esses indicadores de qualidade e possíveis medidas aplicadas pelo MEC, a partir dos mesmos, devem buscar internamente gerar avaliações que se assemelhem às aplicadas pelo governo, com a finalidade de corrigir possíveis falhas no ensino. Essas avaliações devem utilizar métodos psicométricos eficazes e técnicas que permitam obter resultados comparáveis aos apresentados no cenário nacional. Nesse sentido, o presente estudo busca avaliar provas de habilidades acadêmicas desenvolvidas em equidade ao ENADE e que geram indicadores de qualidade interna de uma IES do interior do estado de São Paulo.

Para que os instrumentos resultem em informações precisas é necessário que sua construção seja pautada em objetivos claros e precisos. Assim, apresentamos no Capítulo 2

os conceitos de habilidades e competências, que são os principais objetivos dos sistemas de avaliação educacional.

O capítulo seguinte traz um método muito utilizado para criação de avaliações de larga escala, método matemático esse conhecido como Teoria de Resposta ao Item (TRI). Na construção dos instrumentos de medida, o método demonstra-se mais preciso do que a Teoria Clássica dos Testes (TCT), ainda muito utilizado em análises psicométricas.

Por fim foi inserido um capítulo que descreve o sistema de avaliação utilizado pela Instituição de Ensino Superior – IES estudada. Sua história e valores considerados pela IES.

Na sequência serão apresentados os resultados obtidos pelo presente estudo, permitindo a conclusão que estudos similares podem trazer contribuições para o processo de avaliação educacional. A pesquisa realizada indica que os dados advindos das aplicações dos instrumentos de avaliação educacional permitem a expansão para novos estudos.

CAPÍTULO 1

SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO NO BRASIL

Na década de 1990, de uma maneira abrangente, o Brasil já investia em avaliações de larga escala nos diferentes níveis de ensino. Para que os resultados dessas avaliações fossem utilizados com segurança e confiabilidade os instrumentos utilizados nesse processo deviam ter validade psicométrica. Como afirmam Souza e Oliveira (2003), todo instrumento de avaliação deve ter boas propriedades psicométricas para ser utilizado em um processo de avaliação.

Para Verhine e Dantas (2005) os resultados de uma avaliação em larga escala, não devem ser um fim por si só, mas seu objetivo é um meio de obtenção de subsídios delineadores no processo avaliativo. Klein e Fontanive (1995) destacam que seu propósito deve ser utilizado como um mecanismo de monitoramento do sistema de educação em todo país.

Um sistema de avaliação deve possuir e organizar informações periódicas do sistema educacional. Atualmente uma técnica muito utilizada é a Teoria de Reposta ao Item (TRI), que dentre suas possibilidades, permite que sejam elaboradas avaliações em que todos os participantes sejam colocados em uma mesma escala, mesmo não tendo respondido os mesmos itens (Klein e Fontanive, 1995). O Brasil utiliza vários processos de avaliação educacional, nos diferentes níveis de ensino, que serão apresentados a seguir.

O primeiro nível de ensino considerado pelo sistema de avaliação brasileiro é o Sistema Nacional de Avaliação da Educação Básica (SAEB) que permite o delineamento de um cenário inicial desse sistema de educação no país. Para o ensino médio existe o Exame Nacional do Ensino Médio (ENEM), atualmente utilizado por diversas Instituições de

Ensino Superior (IES) em substituição ao vestibular. Quanto ao ensino superior, sua avaliação é realizada pelo Exame Nacional de Desempenho do Estudante (ENADE), foco desse trabalho.

Bonamino, Coscarelli e Franco (2002) relatam que existe polêmica na definição das avaliações em larga escala promovidas pelo Brasil. Os autores resumem que o tema é cercado por iniciativas de melhoria na qualidade da educação, do papel desempenhado pela avaliação no acompanhamento das políticas educacionais e da associação entre avaliação e promoção de políticas de equidade.

Relatando de maneira sucinta essas avaliações, a educação básica, utiliza o instrumento denominado SAEB (Sistema Nacional da Educação Básica), aplicado pela primeira vez em 1990, tem a concepção de gerar informações que permitam a elaboração de um cenário nacional, que abranja as diversas realidades dos sistemas estaduais e municipais de ensino. Destaca-se que alguns objetivos desse instrumento são os de monitorar a qualidade do sistema de educação básica; oferecer às administrações públicas de educação informações técnicas e gerenciais que lhes permitam formular e avaliar programas de melhorias; e, por fim proporcionar aos agentes educacionais e à sociedade uma visão clara e concreta dos resultados dos processos de ensino e das condições que são desenvolvidos (Castro, 2000).

O ENEM, outro instrumento de avaliação em larga escala utilizado pelo governo brasileiro tem o principal objetivo de aferir o desenvolvimento de habilidades e competências esperado do aluno ao final da educação básica, permitindo, assim, a mensuração de parâmetros para continuidade de seus estudos ou absorção pelo mercado de trabalho (Castro, 2000, Teixeira, 2002). As competências avaliadas pelo ENEM divulgadas pelo INEP são descritas como: dominar linguagens; compreender fenômenos; enfrentar

situações-problemas; construir argumentações e elaborar propostas, esse artigo traz para discussão o conceito de habilidade e competência tratado de maneira diferenciada na concepção das questões do exame.

A seguir, ainda considerando o contexto de avaliação educacional, a avaliação no ensino superior será abordado de maneira mais detalhada por se tratar do foco principal deste trabalho.

A partir da década de 1980, conforme apresenta o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira do Ministério da Educação (MEC/INEP, 2007), ocorreram importantes mudanças na Educação Superior que abrangeram os países industrializados e latino-americanos. Segundo esse instituto, o mercado educacional globalizado, a reforma do nível de ensino se tornando mais dinâmico, a maior diversificação dos tipos de instituições e dos perfis dos docentes, permitiram a ampliação das matrículas e conseqüentemente o aumento da demanda e principalmente da competitividade. Por outro lado, esta globalização educacional e a internacionalização do conhecimento trazem consigo o desafio de proporcionar uma educação superior de qualidade com inovações que permitam diminuir as assimetrias sociais (MEC/INEP, 2007).

Perante este cenário de mudanças sociais e econômicas, o documento do MEC/INEP (2007) ressalta que o processo de avaliação e regulação da educação superior no país tem papel de destaque. O documento ressalta ainda que este processo tem como base diversos argumentos, tais como a necessidade dos estados assegurarem a qualidade e os controles regulatórios, a distribuição e o uso adequado dos recursos públicos; e que, a função atribuída pelo estado à educação superior no País é determinante da proposta de avaliação, tendo de um lado o modelo de inspiração anglo-americana que se baseia predominantemente em gerar resultados quantitativos para produzir resultados

classificatórios e de outro lado os modelos Holandês e Francês, que combina dimensões quantitativas e qualitativas com ênfase na avaliação institucional e análise (INEP, 2007**Erro! Indicador não definido.**).

Conforme descrito, a avaliação de larga escala permite conhecer melhor a dinâmica dos processos, possibilitando que seus resultados tornem-se questões estratégicas das políticas para diversos países, com distintas culturas e orientações ideológicas. O sistema nacional de avaliação está presente em países como Cuba, Chile, México, Peru, China, Espanha, Inglaterra, Austrália, EUA, França, entre outros. Na América Latina, em 19 países existe um sistema nacional de avaliação e crescente participação em avaliações internacionais, como o Programa Internacional de Avaliação do Estudante (PISA), o Laboratório de Avaliação da Qualidade da Educação da UNESCO (LECE) e a Avaliação de Matemática e Ciências (TIMMS), como também ocorrem em países da União Europeia, América do Norte, Ásia e África (Castro, 2009).

O Brasil possui o sistema de avaliação educacional delineado pelo modelo de inspiração inglesa, esse modelo remete à questão da eficiência ou ineficiência das instituições em se adaptarem as exigências sociais, entendendo que a educação superior funciona como fator de incremento no mercado de trabalho. Nesse modelo, a avaliação se realiza como atividade predominantemente técnica, que busca mensurar os resultados produzidos pelas instituições em termo de ensino e também da pesquisa e prestação de serviços à comunidade. Sua ênfase recai sobre indicadores quantitativos que promovem um balanço das dimensões mais visíveis e facilmente descritíveis, a respeito das medidas físicas, como área construída, titulação dos professores, descrição do corpo docente, discente e servidores, relação dos servidores, dos produtos, das formaturas, volumes de

insumos e no caso de citações numéricas estabelecendo muitas vezes *rankings* de instituições (INEP, 2007).

A história da avaliação educacional no Brasil começa com a mais antiga e duradoura experiência bem sucedida na educação superior que são as dos cursos e programas de pós-graduação, desenvolvidas desde 1976 pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (INEP, 2007). Nas duas últimas décadas, discussões e ações relativas à obtenção de resultados dos cursos de graduação levaram à implantação de diferentes propostas de avaliação da Educação Superior (Castro, 2009).

Perante esse cenário, o Programa de Avaliação da Reforma Universitária – PARU criado em 1983, foi a primeira proposta de avaliação da Educação Superior no país. O PARU elaborou questionários que foram respondidos por estudantes, dirigentes universitários e docentes. Tratando basicamente de dois temas: gestão e produção/disseminação de conhecimento, utilizando-se de levantamentos e análises de dados institucionais obtidos através de roteiros e questionários preenchidos por estudantes, professores e administradores (INEP, 2007). Este programa foi desativado no ano seguinte, tendo em seu lugar instituídas várias outras iniciativas governamentais, como a constituição da Comissão de Notáveis em 1985 e do Grupo Executivo da Reforma da Educação Superior – GERES em 1986 (Polidori, Marinho-Araujo & Barreyro, 2006).

As IES entre o final da década de 1980 e início da década de 1990 criaram processos de autoavaliação, que buscavam um meio de comunicação entre o Ministério da Educação (MEC) e as instituições federais, representadas pela Associação das Instituições

Federais do Ensino Superior (ANDIFES) que contribuíram para criação do Programa de Avaliação Institucional das Universidades Brasileiras (PAIUB). Esse novo modelo de avaliação foi sustentado no princípio da adesão voluntária das universidades, o programa concebia a autoavaliação como etapa inicial de um processo que, uma vez desencadeado, se estendia a toda a instituição e se completava com a avaliação externa. Embora desativado pelo MEC, a proposta do PAIUB foi concebido seguindo os princípios de globalidade; comparabilidade; respeito à identidade institucional; não premiação ou punição; adesão voluntária; legitimidade e continuidade do processo avaliativo. Para atingir seus objetivos, o programa foi dividido em quatro etapas, onde a primeira etapa referia-se ao diagnóstico, construído a partir de dados quantitativos sobre o curso em análise; a segunda etapa incluía a avaliação interna, referente à autoavaliação realizada pela comunidade acadêmica do curso; sua terceira etapa contemplava a avaliação externa, realizada pela comunidade acadêmica de diversas IES, profissionais de áreas, representantes de entidades científicas e profissionais, e empregadores, entre outros; por fim a quarta etapa era a reavaliação interna, realizada a partir dos resultados avaliativos obtidos, permitindo uma análise com a comunidade (Polidori, Marinho-Araujo & Barreyro, 2006).

Com a experiência do PAIUB, ficou estabelecido uma nova forma de relacionamento com o conhecimento e a formação, consolidando um diálogo entre a comunidade acadêmica e a sociedade. Mesmo que sua existência tenha sido curta, conseguiu dar legitimidade à cultura da avaliação e proporcionou mudanças visíveis na dinâmica universitária. Embora, tenha recebido ampla adesão das universidades brasileiras, sua implementação foi afetada pela interrupção do apoio do MEC, o que o transformou em

um instrumento interno de avaliação e seu ritmo de desenvolvimento sofreu impactos negativos (INEP, 2007).

A partir da Lei número 9.131/1995 (Brasil, 1995) e da Lei de Diretrizes e Bases da Educação Nacional – Lei número 9.394/1996 (Brasil, 1996) – foram progressivamente implementados novos mecanismos de avaliação: o Exame Nacional de Cursos - ENC, realizado por concluintes de cursos de graduação; o Questionário sobre condições socioeconômicas do aluno e suas opiniões sobre as condições de ensino do curso frequentado; a análise das condições de ensino; a avaliação das condições de oferta; e a avaliação institucional dos Centros Universitários. Cabe destacar que, enquanto no PAIUB a preocupação estava com a totalidade, com o processo e com a missão da instituição na sociedade, no ENC a ênfase é sobre o resultado, com a produtividade, a eficiência, com o controle do desempenho frente a um padrão estabelecido e com a prestação de contas. O PAIUB tem como referência a globalidade institucional, que compreende todas as dimensões e funções da instituição de ensino. O ENC tem como foco o curso, em sua dimensão de ensino, e tem função classificatória, com vistas a construir bases para uma possível fiscalização, regulação e controle por parte do estado, baseada na lógica que a qualidade de um curso é igual a qualidade de seus alunos (INEP, 2007).

A primeira aplicação do ENC ocorreu em novembro de 1996, sendo os cursos de Administração, Engenharia Civil e Direito os pioneiros neste processo de avaliação. Sua aplicação ocorreu no mês de junho dos anos subsequentes e em 2003 o último ano de sua aplicação, totalizando vinte e seis áreas avaliadas. A realização do ENC ficou estabelecida como condição prévia à obtenção do diploma, sendo registrado no histórico escolar do aluno apenas a data de sua efetiva participação (Paiva, 2008). O autor salienta ainda que

dois princípios são considerados, mesmo que indiretamente, pelo ENC, a saber, equidade e obrigatoriedade. Mesmo sendo aplicado apenas para os alunos em fase de conclusão de curso o processo de equidade foi adotado como princípio para sua elaboração. E a obrigatoriedade foi garantida pela impossibilidade de dispensa, legalmente amparada, e pela restrição a obtenção do diploma sem a comprovação de participação no exame.

Durante o período de sua execução o ENC movimentou a comunidade acadêmica, a mídia e a sociedade, grandes consumidores de informações por ele gerados. Inevitavelmente ocorreram acertos e tropeços inerentes a qualquer experiência inovadora. Ao mesmo tempo em que foi aplaudido e reconhecido pelas instituições de ensino superior (IES) com cursos no topo da escala de avaliação, outras vezes foi combatido e repudiado por aqueles cujos resultados estavam no extremo oposto da mesma escala de avaliação. É incontestável que o ENC despertou atenção e interesse da mídia para a questão da qualidade da educação superior brasileira. Colaborando ainda diante da sociedade e dos estudantes de graduação e ensino médio, mesmo que às vezes erroneamente revelações importantes quanto ao desempenho dos concluintes dos cursos de graduação acerca de questões de currículo do curso e da versatilidade para vencer desafios do exercício profissional. Trouxe também, importantes elementos para reflexão e identificação de falhas no projeto pedagógico de curso e sua execução, visando a melhoria de sua qualidade do curso e conseqüentemente alcance de melhor desempenho no ENC (Paiva, 2008).

Nas primeiras edições do ENC, os resultados eram interpretados segundo ordenação dos desempenhos – média geral dos graduandos do curso – a partir da qual eram determinados cinco grupos, sendo prefixado o percentual de integrantes de cada um dos grupos, isto é, aos 12% de cursos com desempenhos mais fracos foi atribuído o conceito *E*,

aos 18% seguintes, o conceito *D*, aos 40% com desempenho médio, o conceito *C*, e os conceitos *B* e *A* foram atribuídos aos 18% e 12% com desempenho mais altos. Desde 2001, o procedimento de conversão dos valores absolutos do ENC em conceitos é baseado na média geral e no desvio-padrão de cada área avaliada, assim é atribuído o conceito *A* aos cursos com desempenho acima de um desvio-padrão (inclusive) da média geral; *B* aos cursos com desempenho entre meio (inclusive) e um desvio-padrão acima da média geral; *C* aos cursos que tiveram seu desempenho no intervalo de meio desvio-padrão em torno – para mais e para menos – da média geral; *D* aos cursos cujo desempenho estiver no intervalo entre um e meio desvio-padrão (inclusive) abaixo da média geral; e *E* aos cursos com desempenho abaixo de um desvio-padrão (inclusive) da média geral (INEP, 2007).

Durante sua existência, pouco mais de sete anos, o ENC contribuiu para o processo avaliativo do país e para os seus defensores o exame tinha como base uma função diagnóstica, proporcionando um levantamento da realidade do ensino no que tange à situação dos graduandos quanto às habilidades e conteúdos avaliados. Mas o ENC recebeu severas críticas, tais como, uma condição de exame geral desarticulado de um conjunto integrado de avaliações com princípios, objetivos, agentes e ações claramente definidos; o fato de exames gerais semelhantes ao ENC terem sua motivação mais fora do que dentro da escola; a sua racionalidade muito mais mercadológica e reguladora do que acadêmica e pedagógica; a desconsideração do perfil do perfil acadêmico do alunado que ingressa em uma IES, tornando inviável a análise do valor agregado pela instituição aos conhecimentos e habilidades dos seus estudantes; a ausência de comparabilidade entre as provas ao longo do tempo, o que compromete seriamente a capacidade de avaliar os êxitos, insucessos e perspectivas dos cursos; os boicotes por parte dos estudantes e a falta de critério para lidar

com provas entregues em branco; a divulgação dos resultados do ENC desvinculados de outros processos avaliativos, atribuindo a ele centralidade no sistema de avaliação e autoridade exclusiva ao comunicar a suposta qualidade dos cursos e a adoção de políticas de premiação e punição de instituições com base em conceitos gerados por um instrumento e por uma metodologia deficiente e, portanto, incapaz de expressar com confiabilidade a qualidade dos cursos (INEP, 2007).

Em 2003 o MEC criou a Comissão Especial de Avaliação da Educação Superior (CEA), com a finalidade de analisar, oferecer subsídios, fazer recomendações, propor critérios e estratégias para a reformulação dos processos de avaliação da educação superior. A CEA contextualizou seu trabalho numa visão abrangente de avaliação, considerando fundamental que os conceitos de integração e participação estivessem presentes nesta proposta, por considerá-los essenciais para a construção de um sistema de avaliação. No mês de setembro do mesmo ano de sua criação, a CEA publicou documento intitulado “SINAES: bases para uma nova proposta de avaliação da educação superior brasileira”, que no ano seguinte foi instituído como Lei Federal sob nº 10.861, de 14 de abril de 2004 (Brasil, 2004) mantendo o início do título e sendo transformado no Sistema Nacional de Avaliação da Educação Superior (SINAES) (Polidori & cols., 2006).

O SINAES fundamenta-se na necessidade de promover a melhoria da qualidade da educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional, efetividade acadêmica e social e, especialmente, o aprofundamento dos compromissos e responsabilidades sociais. Tem como objetivo assegurar o processo de avaliação das IES, dos cursos de graduação e do desempenho acadêmico de seus

estudantes. Assegurando assim que os três componentes principais sejam objetos de avaliação, descritos como: IES, curso e desempenho acadêmico (Polidori & cols., 2006).

Em seus princípios, o SINAES reconhece a diversidade do sistema de educação superior do país; respeita a identidade, a missão e a história das instituições; entende que as instituições devem ser avaliadas globalmente, ou seja, a partir de um conjunto significativo de indicadores de qualidade vistos em sua relação orgânica e não isoladamente, e ainda busca a continuidade do processo avaliativo. Com a composição dos três eixos apresentados: avaliação da IES, do curso de graduação e do estudante, descrito este último como Exame Nacional de Desempenho dos Estudantes (ENADE).

De acordo com a legislação vigente, no ENADE, a avaliação do desenvolvimento de conhecimentos, habilidades, saberes e competências, ao longo da trajetória vivenciada em cada curso, subsidia-se nas Diretrizes Curriculares, nas oportunidades de articulação teórica e prática, e no modo como as competências foram sendo construídas, em relações partilhadas e dos contextos vivenciados. A opção avaliativa do ENADE reflete a tendência de privilegiar o desenvolvimento de competências entre os aspectos a serem avaliados (Polidori & cols., 2006).

Diferente do ENC, o ENADE é aplicado aos alunos que se encontram no final do primeiro ano (ingressantes) e do último ano (concluintes) de formação dos cursos de graduação. Esta mudança destaca a principal diferença dos dois instrumentos, sendo o primeiro criticado por avaliar somente os estudantes concluintes e não mensurar o conhecimento adquirido no decorrer do curso, o ENADE por sua vez teve em sua concepção a aplicação aos estudantes ingressante e concluintes, permitindo que com os

resultados obtidos a determinação do valor agregado, ou seja, qual a situação que a IES recebeu o aluno e o que foi acrescentado de conhecimento ao estudante na conclusão do curso. De acordo com a legislação vigente par ao exame, anualmente o INEP publica uma portaria estabelecendo as áreas que participarão do exame. A prova possui características bastante diferenciadas do que aquelas costumeiramente trabalhadas com os alunos, bem como nos processos avaliativos que ocorrem durante a construção da vida escolar e acadêmica de cada indivíduo, pois a ênfase recai sobre as expectativas em relação ao perfil profissional que se deseja formar em cada curso.

Com a mudança da legislação até o ano de 2009 o INEP utilizava processos amostrais para aplicação da prova. De acordo com a Portaria Normativa a Número 1 (Brasil, 2009) os estudantes ingressantes, que possuíam um porcentual de 7 a 22% de sua carga horária concluída do curso, bem como os estudantes concluintes, aqueles cuja carga horária concluída era superior ou igual a 80% deveriam ser inscritos para participar da prova, mas o INEP aplicaria um processo de amostragem simples para selecionar os estudantes que deveriam participar da prova.

Nova mudança ocorreu em 2010, onde o processo de amostragem simples foi retirado e assim todos os estudantes inscritos realizariam a prova. Os critérios para identificação dos alunos ingressantes e concluintes seguiam aos mesmos aplicados no ano de 2009 (Brasil, 2010).

Para o ano de 2011, ocorreu a mudança mais significativa, onde apenas os estudantes concluintes participaram da prova. Os alunos ingressantes ainda foram inseridos, mas estavam dispensados de participar da avaliação. O MEC justificou a alteração

informando que as notas atribuídas aos ingressantes seriam advindas de outro instrumento de avaliação, que iria permitir mensurar melhor o valor agregado do aluno, sendo esse instrumento o ENEM (Brasil, 2011).

A prova do ENADE é constituída por questões de natureza objetiva e discursiva, que valorizam temas contextualizados e atuais, problematizados em forma de estudo de caso, situações problema, simulacros e outros, proporcionando respostas por meio de múltipla escolha. Ainda sobre a prova, é importante destacar que é dividida em duas partes: a primeira denomina-se Formação Geral (FG), um componente comum às provas das diferentes áreas aplicada a todos os cursos que participam do ENADE. O objetivo é investigar competências, habilidades e conhecimentos gerais que os estudantes já tenham desenvolvido na sua vida; e a segunda parte denomina-se Componente Específico (CE), e contempla a especialidade de cada curso, tanto no domínio dos conhecimentos quanto nas habilidades esperadas para o perfil profissional. Todas as questões das duas partes da prova são elaboradas em níveis diferenciados de complexidade, permitindo que o estudante de 1º ano, responda a várias questões, minimizando o risco de escolha de respostas a apenas algumas questões (Polidori & cols., 2006).

Novas diretrizes do sistema de avaliação superior foram apresentadas por meio da Portaria Normativa 40 (Brasil, 2011b), onde os conceitos atribuídos ao ENADE passaram a ter um peso maior na geração de um indicador de qualidade, que indica a necessidade de uma visita in loco ou renovação de reconhecimento do curso participante do exame. Esse indicador se denomina Conceito Preliminar de Curso (CPC) que é o resultado da nota do ENADE e demais insumos constantes das bases de dados do MEC. O cálculo do CPC é calculado no ano seguinte ao da realização do ENADE de cada área, e se baseia na

avaliação de desempenho de estudantes, corpo docente, infraestrutura, recursos didático-pedagógicos e demais insumos, seguindo orientações técnicas aprovadas pela Comissão Nacional de Avaliação da Educação Superior (CONAES).

Instituída pela Lei 10861 de 14 de abril de 2004 (Brasil, 2004) a CONAES tem como responsabilidade gerir todo processo de avaliação superior vinculado ao MEC. Suas principais funções são: (i) propor e avaliar as dinâmicas, procedimentos e mecanismos da avaliação institucional, de cursos e de desempenho dos estudantes; (ii) estabelecer diretrizes para organização e designação de comissões de avaliação, analisar relatórios, elaborar pareceres e encaminhar recomendações às instâncias competentes; (iii) formular propostas para o desenvolvimento das instituições de educação superior, com base nas análises e recomendações produzidas nos processos de avaliação; (iv) articular-se com os sistemas estaduais de ensino, visando a estabelecer ações e critérios comuns de avaliação e supervisão da educação superior; (v) submeter anualmente à aprovação do Ministro de Estado da Educação a relação dos cursos a cujos estudantes será aplicado o Enade; (vi) elaborar o seu regimento, a ser aprovado em ato do Ministro de Estado da Educação; (vii) realizar reuniões ordinárias mensais e extraordinárias, sempre que convocadas pelo Ministro de Estado da Educação.

O panorama demonstrado sobre o sistema de avaliação no Brasil permite a conclusão que o processo avaliativo ainda está em processo de amadurecimento. Começando pelo SAEB e terminando no ENADE, buscando a experiência das avaliações da CAPES, as análises psicométricas dos resultados obtidos permitem um delineamento mais preciso dos escores seus padrões de qualidade.

O presente estudo procurou trazer os diversos instrumentos disponíveis no sistema de avaliação da educação no país, mas o foco principal é o ENADE. Portanto, a Educação Superior na área de computação será abordada a seguir, começando pelas definições de habilidades e competências.

CAPÍTULO 2

HABILIDADES E COMPETÊNCIAS AVALIADAS PELOS EXAMES NACIONAIS

A elaboração de sistemas padronizados de avaliação de larga escala tende a avaliar o produto da aprendizagem. Assim, se bem construídas poderão, também, permitir a discussão dos diferentes momentos dos processos de construção do conhecimento. Nesse contexto, começamos com alguns termos principais quando se utiliza instrumentos de avaliação educacional.

Começamos pela habilidade, termo utilizado por Carroll como uma forma mais ampla e generalista, sendo sua principal função determinar as variações individuais do potencial dos sujeitos na realização de determinadas tarefas e conteúdos, relata Primi e cols. (2001). E Vendramini (2000) ressalta a importância para o uso de palavras específicas e conceitos claramente definidos ao utilizar-se de concepções alternativas dos conceitos de habilidade e competência.

Carroll (1993) refere-se ao conceito habilidade cognitiva como um destaque individual ou um diferencial particular, que o autor caracteriza como um traço latente que se apresenta de maneira diferente em determinados conjuntos de tarefas. O modelo psicométrico utilizado pelo autor permite aferir determinadas capacidades em um determinado tipo de tarefa que deve atender a alguns padrões básicos e aceitáveis, tais como confiabilidade e homogeneidade (Carroll, 1992).

Por meio de diversas pesquisas em artigos científicos, Carroll (1992) destacou 461 conjuntos de dados, relevantes à temática habilidades cognitivas, que permitiram concluir que sua estrutura pode ser descrita por meio de um modelo de três camadas ou estratos. O

método utilizado pelo autor foi o de análise fatorial exploratória cujos resultados produziram a Teoria dos Três Estratos, cujas habilidades mentais ou inteligências são relacionadas nessas três camadas hierarquicamente distribuídas.

A referida teoria tem como Estrato III a definição do fator g, ou seja, o fator mais abrangente. Já o Estrato II, é onde se localiza as habilidades gerais, composto por dez fatores gerais que influenciam uma quantia relevante de capacidades individuais, a saber, inteligências fluida e cristalizada; memória de curto prazo, leitura e escrita, processamento visual, entre outras. E o Estrato I é composto por um conjunto de habilidades específicas, descritas como capacidade de visualização, conhecimento léxico, habilidade fonológica e diversos outros referentes ao tempo de reação (Carroll 1992, 1993).

A partir de estudos realizados por R. B. Cattell na década de 40, por J. L. Horn na década de 60 e por J. B. Carroll na década de 90, foram observadas semelhanças entre as teorias defendidas por esses autores, o que levou à junção das mesmas e dessa forma surge a Teoria Cattell-Horn-Carroll (CHC) das Habilidades Cognitivas. A representação gráfica da estrutura hierárquica das habilidades cognitivas identificada por Cattell-Horn-Carroll, conhecido como modelo CHC é apresentada pelo esquema adaptado na Figura 1, cujo Estrato III define o fator geral da inteligência (G). O Estrato II é referente aos dez fatores ligados às áreas do funcionamento cognitivo, a inteligência fluida (Gf), a inteligência cristalizada (Gc), o conhecimento quantitativo (Gq), a leitura e escrita (Grw), a memória de 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 de decisão/reação (Gt). Para o Estrato I, encontram-se cerca de setenta fatores específicos que se relacionam com o conjunto de

tarefas normalmente avaliados por testes em geral (Carroll, 1993; McGrew, 2004; Schelini, 2006).

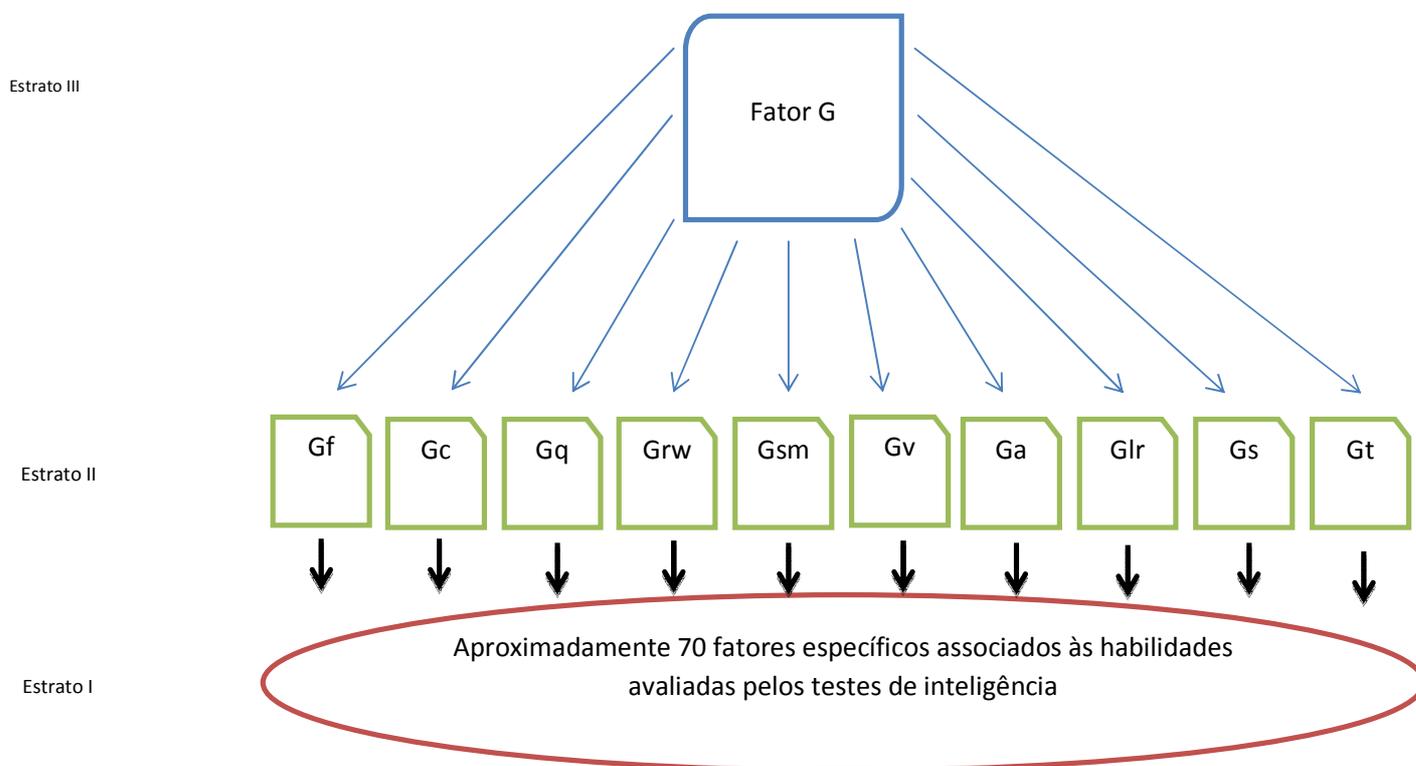


Figura 1. Estrutura hierárquica das habilidades cognitivas de Cattell-Horn-Carroll adaptada de McGrew (2004).

As pesquisas de Carroll foram de certa forma a primeira ideia de classificação com base empírica, dos indicadores das habilidades cognitivas. Destaca-se que o modelo CHC, em 2008, ainda não se encontrava finalizado, mas sim era um parâmetro inicial para novas investigações psicométricas das capacidades humanas (McGrew, 2008).

De forma resumida, o modelo proposto por Carroll não foi a primeira teoria relacionada às habilidades cognitivas. Outros estudiosos como Spearman e Thurstone também apresentaram métodos que buscavam a definição de maneira estruturada para as capacidades mentais, bem como os fatores que determinavam o processo do pensamento,

com base em modelos fatoriais da inteligência (Almeida, 1994; Ribeiro, 1998). Autores como Hair, Anderson, Tetham e Black (2005) caracterizaram tais modelos como oriundos de um conjunto de técnicas estatísticas e análise fatorial, que permitiam a identificação de fatores originários das correlações entre variáveis observáveis.

Para McGrew (2004) a inteligência fluida (Gf) refere-se aos componentes não verbais, com pouca dependência de aspectos culturais, mas com influência maior de aspectos biológicos. Assim, pode-se dizer que, a Gf é derivada da capacidade de raciocínio que implica na capacidade de estabelecer relações entre as ideias, na geração de conceitos abstratos e conseqüentemente na derivação de conclusões lógicas a partir de determinadas situações, o qual seu conhecimento não foi memorizado anteriormente.

Para inteligência cristalizada (Gc) sua associação está ligada a capacidades de solução dos problemas cotidianos. Seu crescimento se dá no cotidiano, com experiências culturais e educacionais, cuja presença é marcante nas atividades acadêmicas, embora não se deva confundir com desempenho escolar, que não tem o mesmo significado (Carroll, 1993; McGrew, 2004; Schelini, 2006). Outros autores como Ackerman e Rolfhus (1999) e Carroll (1993), entendem que a inteligência cristalizada de acordo com o aumento da faixa etária se torna maior, em contrapartida a situação é diferente para inteligência fluida que acaba diminuindo, motivada pela degeneração natural das estruturas fisiológicas.

O modelo de Cattell-Horn é considerado por Carroll (1993) como o mais bem fundamentado e que se aproxima a uma teoria aceitável da estrutura cognitiva das habilidades. Por esse pressuposto, os resultados das pesquisas realizadas por Cattell, Horn e Carroll, baseia-se nesse modelo e pôde aperfeiçoá-lo com o surgimento da Teoria dos Três

estratos da Inteligência, conforme citado anteriormente. Dessa forma, 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* (continua)

Estrato II	Descrição	Habilidades Específicas – Estrato I
<i>Gf</i> – Inteligência Fluída	Operações mentais de raciocínio associadas à capacidade de resolver problemas que impedem 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 escuta, 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.

Tabela 1. *Descrição dos dez fatores amplos do modelo CHC e habilidades relacionadas* (conclusão)

Estrato II	Descrição	Habilidades Específicas – Estrato I
<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, resistência à distorção de estímulos auditivos, memória para padrões sonoros, codificação temporal, discriminação e julgamento musical, manutenção e julgamento de ritmos, intensidade sonora, frequência de discriminação sonora, escuta e discurso de fatores liminares, padrões absolutos e localização de sons.
<i>Glr</i> – Armazenamento e Recuperação da Memória a Longo Prazo	Capacidade associada à extensão e fluência em que as informações ou conceitos são recuperados pela memória de longo prazo por 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.	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 à estímulos mais simples.	Tempo de reação, tempo de escolha, velocidade.

Fonte: ENADE: Análise de itens de Formação Geral e de Estatística pela TRI, de Nogueira (2008, p. 16)

Ainda sobre o modelo CHC, McGrew (2008) descreve um espaço existente no modelo, que é complementado com a inclusão de mais seis grandes fatores referente às habilidades cognitivas, assim permite a expansão do Estrato II para dezesseis categorias de habilidades. Desse modo, foram acrescentadas as seguintes habilidades, Conhecimento Específico (Gkn), Habilidade Tátil (Gh), Habilidade Sinestésica (Gk), Habilidade Olfativa (Go), Habilidade Psicomotora (Gp) e por fim, Velocidade Psicomotora (Gps).

As habilidades cognitivas que representam o desempenho acadêmico por Flanagan, McGrew e Ortiz (2000) são o fator Gq (conhecimento quantitativo), o fator Grw (leitura e

escrita), o fator Gf (inteligência fluida) e o fator Gc (inteligência cristalizada). Os autores destacam que o fator que mais contribuem para o desenvolvimento das habilidades acadêmicas é a Inteligência Cristalizada, pois é um fator intermediário entre o processo do pensamento, o desenvolvimento da linguagem e os processos que determinam o aprendizado.

Revisando a literatura, Ackerman e Rolfhus (1999) relatam que alguns pesquisadores atribuem o desempenho satisfatório em determinadas tarefas e nas tomadas de decisão aos fatores Gc e Gf, isso resume de maneira simples a capacidade intelectual dos adultos. Nesse pressuposto os autores sugeriram uma teoria da inteligência que insere o domínio do conhecimento específico como parte da inteligência em adultos. Esse conhecimento específico assume papel importante no desempenho intelectual e pode ser atribuído aos níveis mais elevados e integrados do conhecimento.

Para Rolfhus e Ackerman (1999) o conhecimento específico está diretamente ligado a Gc. Uma característica que distingue a Gc e Gkn para Ackerman e Lohman (2006) e McGrew e Evans (2004), está no fato de que a Gc representa principalmente o conhecimento geral acumulado de maneira empírica pelo indivíduo em determinada situação, já a Gkn é representada pela amplitude e profundidade do conhecimento em domínios específicos. Dessa forma, o fator Gkn representa melhor a questão do bom desempenho em provas ou testes acadêmicos.

De uma maneira resumida sobre as teorias da inteligência, começando pelo conceito do fator g de Spearman, seguido dos sete fatores da inteligência proposto por Thurstone, dos estudos de Cattell que culminaram na teoria Gc e Gf, acrescentado por Horn os dez fatores

responsáveis pelas habilidades cognitivas, que posteriormente geraram o modelo CHC de Carroll e por fim, a inclusão de mais seis habilidades cognitivas apresentadas por McGrew (2008) que tais estudiosos estabeleceram parâmetros para mensurar o construto em questão. Dessa maneira o construto habilidade acadêmica pôde ser mais bem identificado, quando se trata de habilidades necessárias para aferir o desempenho escolar.

Para Hambrick (2005) as atitudes e interesse de um indivíduo aumentam seu conhecimento por meio do envolvimento em atividades relevantes. Portanto, quanto mais a pessoa sabe sobre um conteúdo específico, fica mais fácil adquirir conhecimentos adicionais sobre aquele assunto.

A habilidade sugere facilidade em trabalhar diversos tipos de informação e para que habilidade se transforme em competência é necessário a existência de experiências anteriores de aprendizagem, caso contrário, não haverá competência (Primi & cols., 2001). Por isso, alguns fatores estão associados à aquisição de competências, ao desenvolvimento de habilidades, à quantidade de investimento determinado pelo sujeito e à qualidade das experiências de aprendizagens (Primi & cols. 2001).

O presente trabalho adota a terminologia de habilidades acadêmicas e competências propostas por Brito (2008), relacionado com o contexto avaliativo do ENADE. Que resumidamente indica a capacidade do indivíduo em realizar determinadas tarefas, solucionar problemas, dominar assuntos específicos, conseguindo realizar uma determinada atividade. Porém, esse aspecto permite que sejam identificados manifestações dos componentes de uma determinada habilidade.

Para Primi e cols. (2001) as instituições educacionais procuram desenvolver as habilidades e competências necessárias para avaliações. Nesse contexto, as instituições de ensino estão se preocupando com o desenvolvimento de competências e a melhora das habilidades dos estudantes, assim permitem uma contribuição para vida acadêmica e por consequência melhor desempenho acadêmico (Souza, 2005).

Outras variáveis podem interferir diretamente no desempenho acadêmico, Soares, Ribeiro e Castro (2001) destacam que o desempenho sofre influência direta do ambiente cultural, do envolvimento da família, de sua situação socioeconômica, bem como características individuais que podem incluir a motivação para o estudo, habilidades cognitivas e desempenho oriundo de níveis de ensino anteriores. Dessa forma os autores indicam que as instituições devem observar esses pontos, pois indicam fatores importantes na aquisição de conhecimento.

Os instrumentos de avaliação buscam mensurar as habilidades acadêmicas, por meio de estudos psicométricos realizados com seus resultados. Os métodos psicométricos utilizados permitem identificar a qualidade e eficiência de um determinado teste utilizado nesse processo de avaliação. E nesse sentido, um dos métodos utilizados para averiguar a qualidade das medidas é a Teoria de Resposta ao Item (TRI), apresentado no próximo capítulo.

Habilidade e Competências Acadêmicas para a área de Computação

Estudo realizado por Castro & Sá (2002) objetivou a descobrir quais as habilidades e competências para o profissional de Computação e Informática, para uma comunidade local no município de Aracaju/SE. Os participantes eram alunos de IES do estado do

Sergipe; profissionais da área formados nestas instituições, entre 1 e 5 anos de formados, e atuantes na área de desenvolvimento de soluções informatizadas, no mesmo município; e gerentes de informática das organizações que empregam estes profissionais, totalizando 138 sujeitos. O instrumento foi um questionário de opinião que buscava caracterizar o perfil sociodemográfico das categorias pesquisadas e o grau de importância de cada atributo do perfil proposto para este profissional. Importante destaca sobre esta pesquisa que a conclusão sobre as categorias estabelecidas pela pesquisa, identificaram habilidades e competências diferentes, mas com alguns aspectos em comum a saber, associação de conhecimento teórico com prático; facilidade de aprendizagem; proposição de soluções; raciocínio lógico; relacionamento humano e trabalho em equipe (Castro & Sá, 2002).

O estudo apresentado acima foi realizado em 2002, três anos antes da aplicação do primeiro ENADE, que ocorreu em 2005. Como os cursos na área de Computação não possuem diretrizes curriculares definidas, no ano de aplicação do primeiro ENADE para o curso de Ciência da Computação, foram diferentes das diretrizes traçadas pelo INEP quando da elaboração da prova. Vale destacar que o ENADE 2005 e 2008 exigiam basicamente as mesmas habilidades e competências, apresentadas a seguir: dominar os fundamentos científicos e tecnológicos relacionados à área de Computação; saber modelar e especificar soluções computacionais para diversos tipos de problemas; ter capacidade para analisar, projetar, desenvolver, implementar, validar e gerenciar qualquer projeto de software; ser apto a projetar e desenvolver sistemas que integrem hardware e software; possuir capacidade para aplicar seus conhecimentos de forma independente e inovadora, acompanhando a evolução do setor e contribuindo na busca de soluções nas diferentes áreas aplicadas; ser empreendedor e ter capacidade de alavancar a geração oportunidades de

negócio na área; conhecer e respeitar os princípios éticos da área de Computação e ter uma visão humanística crítica e consistente sobre o impacto de sua atuação profissional na sociedade (INEP, 2005; INEP 2008).

Pelo ciclo do SINAES, no ano de 2011 a área de Computação foi novamente avaliada e novas definições de habilidades e competências para o curso de Ciência da Computação foram divulgadas. Ocorreu uma evolução natural do processo de sua concepção, o que tornou as habilidade e competências mais precisas e focadas, descritas como: compreender os fatos essenciais, os conceitos, os princípios e as teorias relacionadas à Ciência da Computação para projetar e desenvolver sistemas computacionais e suas aplicações; reconhecer a importância do pensamento computacional no cotidiano e sua aplicação em circunstâncias apropriadas e em domínios diversos; identificar e gerenciar os riscos que podem estar envolvidos na operação de equipamentos de computação (incluindo os aspectos de confiabilidade e segurança); identificar e analisar requisitos e especificações para problemas específicos e planejar estratégias para suas soluções; especificar, projetar, implementar, manter e avaliar sistemas de computação, empregando teorias, práticas e ferramentas adequadas; conceber soluções computacionais a partir de decisões visando o equilíbrio de todos os fatores e restrições envolvidas; empregar metodologias que visem garantir critérios de qualidade ao longo de todas as etapas de desenvolvimento de uma solução computacional; analisar quanto um sistema baseado em computadores atende aos critérios definidos para seu uso corrente e futuro (adequabilidade); gerenciar projetos de desenvolvimento de sistemas computacionais; aplicar temas e princípios recorrentes, como abstração, complexidade, principio de localidade de referência (caching), compartilhamento de recursos, segurança, concorrência, evolução de sistemas, entre outros, e reconhecer que

esses temas e princípios são fundamentais à área de Ciência de Computação; escolher e aplicar boas práticas e técnicas que conduzam ao raciocínio rigoroso no planejamento, na execução e no acompanhamento, na medição e gerenciamento geral da qualidade de sistemas computacionais; aplicar os princípios de gerência, organização e recuperação da informação de vários tipos, incluindo texto, imagem som e vídeo; aplicar os princípios de interação humano-computador para avaliar e construir uma grande variedade de produtos incluindo interface do usuário, páginas WEB, sistemas multimídias e sistemas móveis (INEP, 2011).

CAPÍTULO 3

A TEORIA DE RESPOSTA AO ITEM NA AVALIAÇÃO EDUCACIONAL

Para comprovar as propriedades psicométricas correspondentes a um exame educacional, dentre os diversos modelos matemáticos e estatísticos que podem ser utilizados pode-se destacar a Teoria de Resposta ao Item (TRI). A TRI, somente nos últimos anos, ganhou visibilidade com participação e utilização em diversas propostas de avaliação educacional, mas sua origem é oriunda dos anos de 1950, como apontam Pasquali e Primi (2007). Contudo, somente no início dos anos de 1980 conseguiu um desenvolvimento maior, motivado pelo avanço da tecnologia da informática e o desenvolvimento de aplicativos computacionais que permitiram a realização de operações matemáticas complexas dessa teoria e, assim, tornar seu uso viável.

Uma definição para TRI pode ser descrita como um conjunto de modelos que procuram representar a probabilidade de um indivíduo dar certa resposta a um item como função dos parâmetros do item e da habilidade do sujeito. Essa relação destaca que quanto maior a habilidade, maior a probabilidade de acerto ao item. Os vários modelos propostos na literatura dependem fundamentalmente de três fatores: (i) da natureza do item – dicotômicos ou não dicotômicos; (ii) do número de populações envolvidas – apenas uma ou mais de uma e (iii) e da quantidade de traços latentes que está sendo medida – apenas um ou mais de um (Andrade, Tavares & Valle, 2000).

Para os idealizadores dessa teoria, esse modelo tem como diferencial a análise de itens que abrangesse construtos e procedimentos que anteriormente não eram possíveis pela Teoria Clássica dos Testes (TCT). Assim a TCT por ser um modelo que depende do objeto

medido, não tem as mesmas vantagens da TRI, mas ainda é o método muito utilizado em estudos psicométricos.

Ainda sobre a TCT, Pasquali e Primi (2007) ressaltam que alguns problemas acompanham essa teoria, por exemplo, dependendo da amostra utilizada pode existir variação quanto à dificuldade do teste, ou seja, um item pode se tornar mais fácil ou mais difícil para uma determinada população. Os autores ainda relatam outros problemas encontrados na TCT, principalmente nas situações que se referem à discriminação do item, fidedignidade do item e por fim, aos erros de medida.

Com as diferenças apresentadas na TCT, a TRI demonstra que possui um modelo representativo mais forte no âmbito das diferenças referentes à teoria clássica. A TRI tem se destacado no panorama psicométrico, motivado por funcionalidades que permitem uma avaliação mais rigorosa e precisa, dessa forma, possibilita uma intervenção eficaz por meio dos resultados obtidos (Muniz, 2010). Nesse contexto, o autor apresenta as principais diferenças entre a TCT e a TRI na Tabela 2.

Tabela 2. *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) dificuldade, (b) discriminação e (c) acerto ao acaso
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).

Observando as informações apresentadas na Tabela 2 sobre as duas teorias, pode-se concluir que um dos fatores diferentes e mais importantes da TRI é referente ao processo de identificação e da relação entre as propriedades dos itens e dos sujeitos, tendo como base modelos estatísticos. Para 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”.

Motivado por essas evidências, que nos últimos anos a TRI vem sendo destaque como um sistema de medição que se faz presente em diversas áreas e saberes e, principalmente na avaliação educacional. Conde e Laros (2007) revelam a utilização pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) desde 2005 em seus sistemas de avaliação; já que, como se trata de avaliação em larga escala, sua característica permite um conhecimento melhor do desempenho dos estudantes e as principais variáveis responsáveis por esse resultado.

Para avaliação de larga escala é comum que uma população seja definida por determinadas características que podem variar, dependendo dos objetivos do estudo e, portanto, podem ou não ser relevantes para a diferenciação de populações. E assim, é pelo próprio processo de amostragem que se identifica quantas populações estão envolvidas.

Por se tratar de um modelo estatístico, é possível calcular por meio da TRI 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). Entretanto, a habilidade latente não pode

ser mensurada diretamente. Por isso, 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), que determinam a resposta correta em um determinado item não são passíveis de observação direta, mas podem ser objeto de projeções por meio de observações do comportamento do sujeito perante os problemas ou situações representadas com esse propósito. Por isso, a TRI considera que as estimativas dos traços latentes são dependentes das respostas dos sujeitos e das propriedades do item (Embretson & Reise, 2000).

Por tanto, a teoria do traço latente indica um modelo que pode ser representado por meio de uma fórmula matemática, a relação aparente entre variáveis hipotéticas e variáveis observadas, que indicam a origem do nome ‘traço latente’. Os itens de um teste, por exemplo, podem ser considerados variáveis observáveis e quando suas características são conhecidas, essas se tornam constantes na equação e, dessa maneira, solucionáveis. Portanto, é 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 Anastasi e Urbina (2000) e Pasquali e Primi (2007), dois postulados básicos da TRI devem ser considerados. O primeiro 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. Essa técnica é conhecida por modelagem latente, em que o θ é a causa e o desempenho o efeito.

O segundo postulado menciona a Curva Característica do Item (CCI) que se caracteriza por uma função matemática monotônica crescente, que permite visualizar a relação entre traços latentes e o desempenho do sujeito no item. Sendo assim, a medida que a habilidade latente aumenta, proporcionalmente, aumentam as chances, ou seja, a probabilidade de acerto ao item. Outro aspecto importante é seu tratamento de fidedignidade e do erro de mensuração através das funções de informação dos itens.

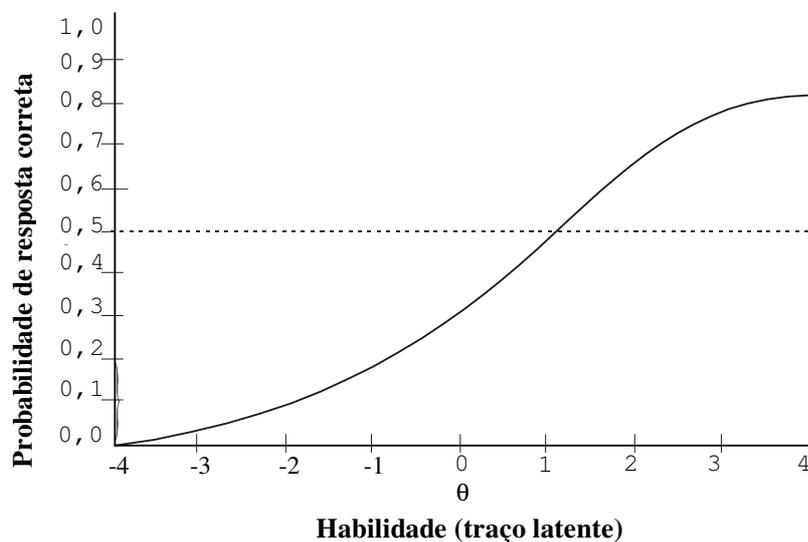


Figura 2. Curva característica do item (CCI)

Os modelos matemáticos são:

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1+e^{D(\theta-b_i)}} \quad , \quad P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \quad E$$

$$P_i(\theta) = c_i + (1-c_i) \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}}$$

Em que:

θ = valor da variável latente

b_i = índice de dificuldade

a_i = índice de discriminação

c_i = probabilidade de acerto ao acaso

$e = 2,72$

$D = 1,7$

É importante destacar ao menos cinco das principais vantagens na utilização da TRI.

A primeira é referente ao cálculo do nível de aptidão do sujeito que pode ser realizada independente do número de itens ou conjunto de itens, desde que meçam o mesmo traço latente, ou seja, a habilidade do sujeito independe do teste. A segunda vantagem é referente ao cálculo dos parâmetros dos itens. Isso indica que os cálculos de dificuldade e discriminação de um item não dependem da amostra utilizada (Hambleton, Swaminathan & Rogers, 1991; Pasquali & Primi, 2007).

Outra vantagem é que teoria permite o emparelhamento dos itens de acordo com a habilidade do sujeito. Assim, podem-se utilizar itens mais difíceis para sujeitos com maior habilidade e itens mais fáceis para sujeitos menos hábeis, resultando em escores que permitam a comparação em ambos os casos. A quarta vantagem é que as suposições não são necessárias, aparentemente, improváveis como, por exemplo, a ideia que os erros de medidas possam ser iguais para todos os sujeitos. E por fim, a última vantagem é que, diferente das exigências da psicometria clássica, a TRI não apresenta a necessidade de se trabalhar com testes exclusivamente paralelos (Hambleton, Swaminathan & Rogers, 1991; Pasquali & Primi, 2007).

Nesse contexto, Andrade e Valle (1998) destacam outra característica, que pode ser considerada como vantajosa na utilização da TRI, que é referente à possibilidade de

equalizar as habilidades dos sujeitos, independente deles pertencerem à mesma população, sendo submetidos ao mesmo item ou conjunto de itens, cujos resultados são passíveis de comparação de seus desempenhos na tarefa solicitada. Portanto, a presente teoria se demonstra como importante ferramenta para a avaliação em larga escala, pois considera a habilidade dos sujeitos em suas respostas a cada item, gerando informações precisas e consolidadas para o contexto educacional.

A TRI muda o foco de análise da prova como um todo para a análise de cada item. Uma de suas importantes propriedades são os fatos dos parâmetros dos itens e as proficiências dos indivíduos serem invariantes. Graças a essa propriedade, a TRI associada a outros procedimentos estatísticos, permite comparar alunos, estimar a distribuição de proficiências da população e subpopulação e ainda monitorar os progressos de um sistema educacional. Um dos procedimentos mais utilizados para colocar indivíduos que fazem provas diferentes em uma mesma escala de proficiência é a utilização de itens comuns nas provas. Dessa maneira é possível comparar o desempenho dos indivíduos (Klein, 2009).

Para Vendramini (2006), a TRI também é utilizada para avaliar mudanças educacionais ou valor agregado. Destaca ainda que uma das vantagens de sua utilização na avaliação educacional é que permite análises qualitativas a partir dos resultados brutos de uma prova, fornecendo assim, informações mais precisas do desempenho dos respondentes e da qualidade dos itens utilizados. E que os itens que compõem a prova devem ter índices de dificuldade e de discriminação aceitáveis e devem estar correlacionadas com a prova total.

Vianna (2003), destaca que as avaliações em larga escala fazem uso da TRI para seleção dos itens a serem aplicados em seus testes. O método já era utilizado em países

desenvolvidos desde os anos sessenta. O autor ainda ressalta que nos estudos sobre as grandes avaliações realizadas no país, independentemente do órgão administrativo que as promova, observam-se alguns problemas complexos e de difícil solução, como os relacionados às escalas empregadas, ao tipo de instrumentação usado e aos julgamentos comparativos que são emitidos sem maiores considerações sobre suas implicações e consequências decorrentes das repercussões no ambiente educacional e suas extrapolações na sociedade. O uso de diferentes tipos de escalas não constitui problema, desde que seus referenciais apresentem pontos comuns que os tornem equivalentes, o que nem sempre ocorre. Assim, os grandes referenciais são quase sempre a média, o desvio padrão e o chamado escore “z”, que expressa a relação da diferença entre o escore obtido e a média do grupo em termos de desvio padrão. Os escores passam a ter valores, teoricamente, entre menos 3,0 e mais 3,0, passando por 0,0, que corresponde à média.

CAPÍTULO 4

UM SISTEMA INSTITUCIONAL DE AVALIAÇÃO

Numa IES do interior de São Paulo foi instituído em agosto de 2008 o Programa de Avaliação Continuada (PAC), oriundo de um sistema de avaliação denominado Prova Pluridisciplinar, que objetivava uma integração maior entre as disciplinas dos currículos de graduação, compondo o sistema de aprovação das disciplinas dos cursos, sua extinção ocorreu também em 2008 devido à alteração no processo de avaliação dos cursos de graduação. O PAC aproveitou a lacuna deixada por seu antecessor e aproveitando dos indicadores gerados pelo ENADE solicitou auxílio do LabAPE – Laboratório de Avaliação Psicológica e Educacional e elaborou-se uma prova piloto para alguns cursos de graduação, utilizando-se de questões âncoras, retiradas das provas do ENADE.

No ano de 2008 os cursos submetidos a avaliação do PAC foram: Administração; Direito e Psicologia, totalizando 2.908 alunos previamente selecionados. Em 2009 a prova PAC ocorreu para os cursos da área de saúde e total de alunos selecionados foi de 2.330, para o ano de 2010 os alunos previamente selecionados representaram um total de 2.380 para os cursos da área de Engenharia, Computação e Pedagogia.

Para o ano de 2011, os cursos avaliados foram os de Administração, Direito e Psicologia, totalizando 2.962 alunos selecionados e no ano de 2012 novamente a área de saúde participou do PAC com um total de 2.229 alunos selecionados.

Em 2011 o antigo PAC se submeteu as adequações naturais de um processo de avaliação e foram implementadas mudanças no processo de conduta e ações inerentes aos seus princípios. Na etapa realizada até do ano de 2010 o objetivo era realizar um

mapeamento da situação dos alunos em relação aos que cursavam o mesmo curso na cidade de oferta, do estado, da região e do país. Resultados que comprovaram informações oficiais divulgadas pelo MEC.

Observando que a primeira etapa foi concluída, entendeu-se necessária a revisão do sistema utilizado pelo PAC. Sendo assim foram discutidas as melhorias e implantadas no ano de 2011, onde o PAC passou a se chamar Programa de Avaliação Discente (PAD), cujo objetivo é avaliar dos alunos do curso e proporcionar uma visão apurada da situação real do curso comparado com os resultados do ENADE e tomar ações para correção de eventuais falhas e conscientizar os alunos da importância da realização do ENADE. Permitindo que o conceito atribuído pelo MEC relate a verdadeira situação em que se encontra o curso avaliado.

EQUALIZAÇÃO OU EQUIPARAÇÃO DE PROVAS EDUCACIONAIS

O conceito equalização de provas é citado por Andrade, Tavares e Valle (2000) como equiparar ou tornar comparável parâmetros de provas distintas ou habilidades de diferentes grupos de respondentes. Esse procedimento é um dos mais importantes da TRI, que permite estabelecer uma métrica padrão, tornando assim uma escala comum de itens ou habilidades comparáveis.

Os autores descrevem que existem seis casos possíveis, quanto ao número de grupos e de tipos de prova envolvidos no processo de equalização, sendo, um único grupo fazendo uma única prova; um único grupo, dividido em dois subgrupos, fazendo duas provas, totalmente distintas (nenhum item comum); um único grupo, dividido em dois subgrupos, fazendo duas provas apenas parcialmente distintas, ou seja, com alguns itens comuns; dois

grupos fazendo uma única prova; dois grupos fazendo duas provas, totalmente distintas (nenhum item comum) e dois grupos fazendo duas provas, apenas parcialmente distintas, ou seja, com alguns itens comuns. A presente pesquisa estudou a equalização por dois grupos fazendo provas parcialmente distintas, ou seja, com alguns itens comuns (Andrade, Tavares e Valle, 2000).

O modelo adotado para equalização, que trata de provas parcialmente diferentes é conhecida como equalização via itens comuns. Neste caso, é possível representar o melhor exemplo do uso e da importância da equalização, ilustrando o maior avanço da TRI sobre a Teoria Clássica. O uso de itens comuns entre provas distintas, permite que todos os parâmetros estejam na mesma escala ao final do processo de estimação, permitindo assim comparações e a construção de escalas do conhecimento interpretáveis, que são de grande importância na área educacional (Andrade, Tavares e Valle, 2000).

Para que exista garantia da qualidade no processo de equalização Silva e Soares (2010) destaca que é importante verificar os seguintes aspectos. A existência de uma simetria entre os escores equalizados; que o grupo ou subgrupo da população determine uma invariância do objeto de estudo e por fim, a invariância no tempo, que os dados de um determinado ano de realização da equalização permitam comparação com as de anos diferentes.

Nunes e Primi (2005), realizaram estudo que buscavam discutir o tamanho mínimo da amostra para utilização da TRI. Os autores estudaram o impacto do tamanho da amostra na calibração dos itens. Para realização da pesquisa, foi utilizada uma base de dados cedida pelo Projeto de Avaliação Externa – ISP/UFBA – FAPEX, contendo as respostas obtidas em uma prova de Avaliação de Desempenho aplicada no ano de 2002. A base de dados

inclui resposta de 44.636 estudantes da 4ª série do ensino básico na disciplina de matemática. Foram criadas, a partir da base de dados original, 9 bases parciais. O instrumento utilizado foi a prova de matemática para Avaliação de Desempenho, composto por 25 itens de múltiplas escolhas, com quatro alternativas. A prova foi elaborada pelo núcleo de conteúdo do Projeto de Avaliação Externa, a partir de matrizes de conteúdos que foram validadas por especialistas em educação, diretores e professores do Estado da Bahia, bem como pela comunidade geral. A prova era composta por itens que nos estudos de pré-testagem haviam atendido às especificações mínimas indicadas pelo setor de psicometria do Projeto de Avaliação Externa.

Os dados foram analisados com a utilização do Software XCalibre, específico para a estimação dos parâmetros psicométricos de itens dicotômicos, por TRI, nos modelos de dois e três parâmetros. Para verificar o impacto do tamanho da amostra na estimativa do nível de dificuldade dos itens, foi calculada a dificuldade da prova a partir da média dos “b” de todos os seus itens. Em seguida, foi calculada a diferença dos valores encontrados nas amostras parciais em relação à amostra completa (com 44 mil estudantes). Também foram calculadas as correlações dos “b” dos itens entre as amostras parciais e a amostra completa.

O estudo concluiu que amostras com 200 sujeitos geram resultados bastante aproximados principalmente quanto aos parâmetros de dificuldade e de capacidade dos sujeitos. Os autores consideram que essa aproximação não é tão eficaz quando se considera os parâmetros de discriminação. No entanto, que mesmo para o parâmetro “a”, a posição relativa dos itens foi estimada, ou seja, os itens que apresentaram maior capacidade de discriminação na amostra com 200 participantes foram os mesmos em amostras maiores.

Silva e cols. (2011) realizaram 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 em relação aos estudantes brasileiros avaliados pelo ENADE. Foram participantes 141.689 estudantes do curso de Administração que participaram do ENADE 2006 e 916 estudantes desse mesmo curso de uma Universidade particular do interior do estado de São Paulo, divididos em quatro campi. 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, pode-se constatar que as questões mais difíceis da prova de conhecimento específico correspondem aos itens 27 (âncora ao ENADE), 18 e 40. Em contrapartida, observou-se que o item 23 corresponde ao mais fácil da prova. Ainda de acordo com o mapa de itens, se pode 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 ajustes *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 evidenciam que o desempenho dos estudantes ingressantes dos quatro campi avaliados foi semelhante ao desempenho dos estudantes ingressantes brasileiros com conceito 2 e 3, conceitos esses instituídos pela matriz do ENADE. 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 apresentaram desempenho entre os Conceitos 2 e 3, ou seja, valores médios mais baixos entre os campi da IES (Silva e 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 74.417 estudantes do curso de Direito divididos em 73.106 estudantes que realizaram o ENADE no ano de 2006 e 1.311 estudantes de uma Universidade particular de ensino superior de dois campi 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 evidenciam 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 estudante no curso (Bartholomeu & col., 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 IES estudada, 80% deles, dos dois campi, foram considerados na faixa de competência baixa. Em presença desses resultados e de discussões acerca do tema, aos autores indicaram que, um dos fatores que podem ter contribuído para os resultados refere-se a falta de motivação na situação de realização da prova.

Aproveitando as possibilidades que o estudo permite, abaixo é descrito os objetivos que definem a importância de pesquisas para temáticas similares.

OBJETIVO GERAL

Analisar uma prova de habilidades acadêmicas equalizada com o Exame Nacional de Desempenho dos Estudantes da área de Computação, por meio da Teoria de Resposta ao Item.

OBJETIVOS ESPECÍFICOS

- Identificar e descrever as propriedades psicométricas (dificuldade, discriminação, precisão e fidedignidade) dos itens, que compõem a prova da área de Computação do ENADE 2005.
- Identificar e descrever o comportamento (displace) dos itens âncoras na prova do PAC 2010.
- Identificar e descrever as propriedades psicométricas (dificuldade, discriminação, precisão e fidedignidade) dos itens que compõem a prova de Ciência da Computação do PAC 2010.
- Verificar as evidências de validade por meio da análise e interpretação do mapa de itens obtidos pela TRI;
- Comparar o desempenho acadêmico entre os alunos que realizaram a prova do PAC 2010 e os que fizeram o ENADE 2005.

CAPÍTULO 5

RESULTADOS E DISCUSSÃO

Para que os objetivos sejam alcançados a pesquisa foi dividida em três estudos, o primeiro teve como proposta analisar e selecionar os itens da prova do ENADE 2005 da área de Computação; e segundo estudo propôs a calibração e equalização dos itens que compuseram a prova PADUSF 2010; e um terceiro estudo que buscou a interpretação do mapa de itens da prova de Ciência da Computação.. A seguir os estudos serão descritos especificamente.

5.1 -Estudo 1 – Análise e Seleção dos Itens da Prova de Computação do ENADE 2005 via TRI

5.1.1 – Método

Participantes

Os participantes desta etapa do estudo pertencem a um banco de dados que contém os resultados dos estudantes da área de Computação, 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. 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.

O banco de dados é composto por estudantes selecionados de acordo com os critérios e rigores do MEC, selecionados aleatoriamente por amostragem estratificada, e contam com informações referentes ao curso, às IES, respostas ao Componente de Formação Geral, ao Componente Específico por curso e ao Questionário Socioeconômico. Os dados do ENADE da área de Computação, realizado em 2005, contem informações de 13.955 alunos de todos os estados brasileiros e distrito federal, sendo que a concentração maior ficou na região sudeste, divididos em 7.221 ingressantes (51,75%) e 6.734 concluintes (48,25%). Na Tabela 3 é apresentado a distribuição dos estudantes de Computação de acordo com a categoria administrativa das IES, demonstrando que as matrículas em Instituições privadas representam a grande maioria (70,3%).

Tabela 3. Distribuição de estudantes no ENADE 2005 de Computação por categoria administrativa

Categoria Administrativa	Total de Estudantes	%
Federal	1.976	14,2
Estadual	1.235	8,8
Municipal	935	6,7
Particular	9.809	70,3
Total	13.955	100

Material

Serviu de material a prova da área de Computação do ENADE realizado em 2005. Cabe ressaltar que a prova do ENADE é composta por 40 questões em formatos discursivas e objetivas de múltipla escolha, sendo que 10 são de avaliação de conhecimentos gerais – a prova de Formação Geral (FG); e, por 30 questões específicas de cada área referentes à parte do Componente Específico (CE). Para o estudo foram utilizadas as questões objetivas, sendo 7 questões de FG e 26 questões de CE,

A prova de FG é comum à 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).

Procedimento

Foram extraídas da prova do ENADE as questões que compuseram a prova do PAC/Computação 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 próximo 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.

Outro critério contempla a medida “*b*”, ou seja, o índice de dificuldade do item. Para compor a prova de Computação do PAC 2010, foram selecionados itens fáceis, medianos e difíceis. E, por último, a questão das cargas dos resíduos completam o último 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.

Foi solicitado à coordenação dos cursos da área de Computação que, com base nos parâmetros psicométricos dos itens do ENADE, selecionassem àqueles que satisfizeram os critérios estabelecidos para este estudo. Após a seleção dos itens que compuseram a prova de Computação 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 dos cursos da área de Ciência da Computação.

5.1.2 – Resultados e Discussão

Para melhor compreensão dos resultados obtidos é importante apresentar de maneira breve alguns aspectos que estavam ligados à aplicação da prova, bem como, informações importantes que permitiram a elaboração do banco de dados utilizado no presente estudo. A prova do ENADE realizada no ano de 2005 com os estudantes do curso de Ciência da

Computação 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, sendo 3 delas do tipo dissertativa, 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 do INEP (<http://portal.inep.gov.br/enade>).

Para o componente de Conhecimento Específico, constituído por 28 questões objetivas e 2 dissertativas, elaborado a partir das Diretrizes Curriculares Nacionais dos Cursos, objetivando levantar indicadores da qualidade do ensino oferecido pelas Instituições de Ensino Superior além de buscar identificar os problemas e as necessidades provenientes da formação do profissional na área de Computação. Esse componente da prova também procurou avaliar o desenvolvimento dos estudantes com relação ao conhecimento profundo de aspectos teóricos da área de Computação (INEP, 2005).

Os resultados do ENADE são apresentados por estatísticas básicas, divulgadas em relatórios específicos a cada curso e disponibilizado no portal do INEP (<http://portal.inep.gov.br/enade>). Na Tabela 4 é apresentado alguns resultados, oriundos do relatório do curso de Ciência da Computação referente ao componente de Formação Geral

do ENADE 2005, em que os resultados apontam para um desempenho geral dos concluintes superior aos dos ingressantes.

Tabela 4. Estatísticas básicas das questões objetivas de Formação Geral por grupo de estudantes

Estatísticas	Total	Grupo	
		Ingressantes	Concluintes
População	57.932	36.324	21.608
Tamanho da amostra	39.374	23.707	15.667
Presentes	32.060	18.004	14.056
Média	54,6	52,4	58,3
Erro-padrão da média	0,1	0,1	0,1
Desvio-padrão	18,7	18,8	18,0
Nota mínima	0,0	0,0	0,0
Mediana	56,9	55,0	61,0
Nota máxima	100,0	100,0	97,0

Fonte: INEP (2005) Relatório síntese

As questões do componente Formação Geral também passaram por tratamento estatístico em que se objetivou dividi-las entre questões 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 consideradas fáceis, os itens deveriam estar na faixa de acerto entre 85% e 61%; da mesma maneira, para serem considerados itens de dificuldade mediana, as questões devem se enquadrar no percentual entre 60% e 41% de acerto; para serem considerados itens difíceis, estes devem apresentar índice de acerto de 16% a 40%; por fim, os itens avaliados como muito difíceis devem ter índice de acerto menor ou igual a 15%. Na Tabela 5 são apresentados os itens correspondentes as questões objetivas do componente Formação Geral, em que se é possível observar que apenas 1 é muito fácil e que a maioria foi considerada fácil. Importante destacar que as tabelas do relatório do INEP, o índice de

dificuldade é denominado como índice de facilidade, mas representa, nos dois casos, o percentual de acertos na questão (INEP, 2005).

Tabela 5. Classificação das questões objetivas de Formação Geral segundo o índice de facilidade

Índice de facilidade	Classificação	Questões
$\geq 0,86$	Muito fácil	4
0,61 a 0,85	Fácil	1, 2, 3, 7
0,41 a 0,60	Médio	5, 6
0,16 a 0,40	Difícil	-
$\leq 0,15$	Muito difícil	-

Fonte: INEP (2005) Relatório síntese

Para o componente de Conhecimento Específico foram realizadas as mesmas análises em que se obtiveram os resultados apresentados na Tabela 6. Conforme as informações contidas nessa tabela, os estudantes concluintes apresentaram desempenho melhor ($M=29,3$) do que o desempenho dos estudantes ingressantes ($M=20,5$). A variabilidade das notas se demonstrou que igualmente ao componente de Formação Geral, para o componente de Conhecimento Específico os concluintes obtiveram um desempenho melhor.

Tabela 6. Estatísticas básicas em Componente Específico por grupo de estudantes

Estatísticas	Total	Grupo	
		Ingressantes	Concluintes
População	25.154	14.575	10.579
Tamanho da amostra	19.991	9.488	7.503
Presentes	13.955	7.221	6.734
Média	24,2	20,5	29,3
Erro-padrão da média	0,1	0,1	0,2
Desvio-padrão	14,2	11,9	15,6
Nota mínima	0,0	0,0	0,0
Mediana	21,8	18,2	27,5
Nota máxima	85,5	77,8	85,5

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 apresentadas na Tabela 7, indicando que a maior parte dos itens pode ser considerada de facilidade de nível médio e os demais enquadrados como difícil e muito difícil (INEP, 2005).

Tabela 7. Classificação das questões objetivas de Conhecimento Específico segundo o índice de facilidade

Índice de facilidade	Classificação	Questões
$\geq 0,86$	Muito fácil	-
0,61 a 0,85	Fácil	-
0,41 a 0,60	Médio	41, 42, 43, 44, 45, 47, 49, 50, 51, 52, 53, 54, 58, 59, 67
0,16 a 0,40	Difícil	60, 61, 62, 63, 64, 65, 66, 68
$\leq 0,15$	Muito difícil	46, 48, 56, 57, 69

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 Computação que participaram do ENADE 2005, bem como de seus perfis profissionais, foi possível a categorização das IES em conceitos. De acordo com o Relatório síntese da área de Computação, esses conceitos são estabelecidos obtendo-se as notas finais dos estudantes ingressantes e concluintes do componente de Formação Geral (com peso de 25% do conceito) e as notas finais da prova de Conhecimento Específico (com peso de 75% do conceito sendo que a nota dos ingressantes tem peso de 15% e dos concluintes peso 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 8 contém os cálculos finais das notas dos estudantes e seus respectivos conceitos.

Tabela 8. Distribuição dos conceitos

Conceito	Notas finais
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 apresentados na Tabela 8 para o cálculo dos conceitos dos cursos de Computação no Brasil, avaliados pelo ENADE no ano de 2005, é possível afirmar, de acordo com a Tabela 9, que a maior parte dos cursos obtiveram o conceito 3, o que representa 41,3% dos cursos participantes. A mesma tabela 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 conceito 3 (INEP, 2005).

Tabela 9. Número e percentual de cursos participantes por grandes regiões segundo o conceito obtido

Conceito	Brasil		Região									
			Norte		Nordeste		Sudeste		Sul		Centro-Oeste	
	N	%	N	%	N	%	N	%	N	%	N	%
1	11	1,6	1	3,1	1	1,3	4	1,1	4	2,7	1	1,5
2	71	10,4	5	15,6	5	6,4	39	10,9	15	10	7	10,6
3	283	41,3	11	34,4	22	28,2	158	44	63	42	29	43,9
4	99	14,5	5	15,6	21	26,9	47	13,1	22	14,7	4	6,1
5	19	2,8	-	-	2	2,6	11	3,1	3	2	3	4,5
SC	202	29,5	10	31,3	27	34,6	100	27,9	43	28,7	22	33,3
Total	685	100	32	100	78	100	359	100	150	100	66	100

Fonte: INEP (2005) Relatório síntese

Análise pela Teoria de Resposta ao Item

Para realização desta etapa do estudo, foi necessário selecionar, junto ao banco de dados fornecido pelo MEC, somente os estudantes do curso de Computação. 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 desse 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 & Write, 1991).

Desta maneira, foi aplicada a TRI nas duas partes da prova do ENADE de Computação 2005 em que, para o componente de Formação Geral obteve-se os resultados apresentados na Tabela 10. De acordo com essa tabela, a questão que solicitou menor habilidade dos respondentes refere-se ao item 4 ($b=-1,62$) 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 se ficam próximos dos resultados divulgados pelo INEP em que a questão 4 obteve uma porcentagem de acerto maior que 86%, avaliada como muito fácil e a questão 6 que obteve uma porcentagem de acerto entre 41% e 60%, sendo o item de nível médio da prova (INEP, 2005).

Tabela 10. Estatísticas dos itens em Formação Geral no ENADE de Computação 2005 calibrados pelo modelo de Rasch

Nº Item FG	Correlação Item total	Índice de Dificuldade Measure	Índice de Ajuste		Resíduo Loading
			<i>Infit</i>	<i>Outfit</i>	
FG 1	0,49	-0,36	0,93	0,87	-0,41
FG 2	0,46	-0,45	0,97	0,96	-0,16
FG 3	0,50	-0,07	0,95	0,92	-0,15
FG 4	0,38	-1,62	0,96	0,83	-0,16
FG 5	0,47	1,17	1,07	1,12	0,97
FG 6	0,48	1,66	1,06	1,15	-0,27
FG 7	0,42	-0,33	1,04	1,04	-0,03

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 superam o valor crítico de 1,40. Contudo, observou-se que a questão de número 6 atingiu na medida de *Outfit* o valor de 1,15, valor bem próximo do ideal, tendo como base que, o item para ser considerado ideal deve obter, nessa medida, o valor de 1,20. Da mesma forma, os índices da medida de 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 1 (-0,41) e a questão 5 (0,97).

Os mesmos procedimentos foram realizados para obtenção da calibragem dos itens pelo modelo de Rasch do Componente Especifico da prova de Computação do ENADE 2005. Os resultados estão distribuídos na Tabela 11 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 47 ($b=-1,50$), mas pertencente ao grupo de itens considerados medianos. Já o item com maior índice de dificuldade foi o de número 46 ($b=0,91$), representando um item muito difícil. Esses resultados também estão de acordo com os resultados divulgados pelo INEP, em que o item 47 obteve um percentagem de acerto 41% a 60%, se encaixando como mediano e o item de número 46, que obteve menor percentagem de acerto ($\leq 15\%$), avaliado como muito difícil (INEP, 2005).

Conforme os dados referentes a Tabela 11 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 56, 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 49 ($r_{item=total}=0,29$), 50

($r_{item=total}=0,23$), 51 ($r_{item=total}=0,26$), 52 ($r_{item=total}=0,26$), 58 ($r_{item=total}=0,20$), 59 ($r_{item=total}=0,29$), 60 ($r_{item=total}=0,24$), 62 ($r_{item=total}=0,28$), 65 ($r_{item=total}=0,25$), 68 ($r_{item=total}=0,23$) e 69 ($r_{item=total}=0,25$), tiveram suas correlações abaixo de 0,30, valor entendido como ideal, totalizando 39,29% 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 11. Estatísticas dos itens no Componente Específico do ENADE de Computação 2005 calibradas (continua)

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	Residual Loading
41	2.852	14040	0,35	1,06	1,12	0,13	-0,27
42	3.862	13998	-0,08	0,95	0,93	0,33	0,30
43	2.972	14.040	0,30	0,91	0,87	0,38	0,33
44	3.928	13.975	-0,10	0,95	0,94	0,34	0,29
45	7.367	13.995	-1,24	0,99	0,99	0,31	0,14
46	1.829	14.021	0,91	1,02	1,06	0,16	-0,08
47	8.269	14.097	-1,50	0,96	0,95	0,35	0,22
48	2.288	13.997	0,64	1,04	1,09	0,15	-0,17
49	5.041	13.970	-0,50	0,99	0,99	0,29	0,08
50	3.515	14.011	0,06	1,01	1,02	0,23	-0,11
51	3.567	13.973	0,04	0,99	1,00	0,26	0,05
52	3.464	13.959	0,08	0,99	0,99	0,26	0,08
53	2.949	13.819	0,28	1,03	1,07	0,19	-0,04
54	2.752	13.828	0,38	1,03	1,07	0,17	-0,23
56	2.302	14.680	0,69	1,09	1,22	0,05	-0,38
57	2.373	14.507	0,63	1,05	1,12	0,12	-0,25
58	3.745	14.558	0,02	1,03	1,04	0,20	-0,13
59	4.133	14.547	-0,12	0,98	0,98	0,29	0,15
60	2.798	14.593	0,43	0,98	1,01	0,24	0,06
61	5.139	14.486	-0,48	0,98	0,97	0,31	0,07
62	4.061	14.563	-0,09	0,98	0,97	0,28	0,04

Tabela 11

Estatísticas dos itens no Componente Específico do ENADE de Computação 2005 calibradas (conclusão)

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-toal	Residual Loading
63	3.846	14.483	-0,02	0,92	0,90	0,38	0,40
64	2.805	14.584	0,42	1,08	1,14	0,10	-0,40
65	3.765	14.423	0,00	1,00	1,00	0,25	-0,26
66	3.133	14.491	0,26	1,06	1,08	0,15	-0,31
67	7.755	14.641	-1,25	0,93	0,91	0,41	0,49
68	3.813	14.485	-0,01	1,01	1,02	0,23	-0,13
69	4.064	14.473	-0,11	1,00	1,00	0,25	-0,08
Média	3.871,0	14.258,5	0,00	1,00	1,02		
DP	1.562,4	287,1	0,56	0,05	0,08		

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 (a esquerda) e o nível de dificuldade dos itens (a direita). Assim, por meio do mapa de itens é possível observar que há uma concentração de estudantes entre o valor -2 a 0. O lado direito do gráfico corresponde à posição dos itens aplicados no ENADE de Computação 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 acima da média para os estudantes avaliados. Os estudantes apresentaram uma habilidade inferior a média de dificuldade dos itens.

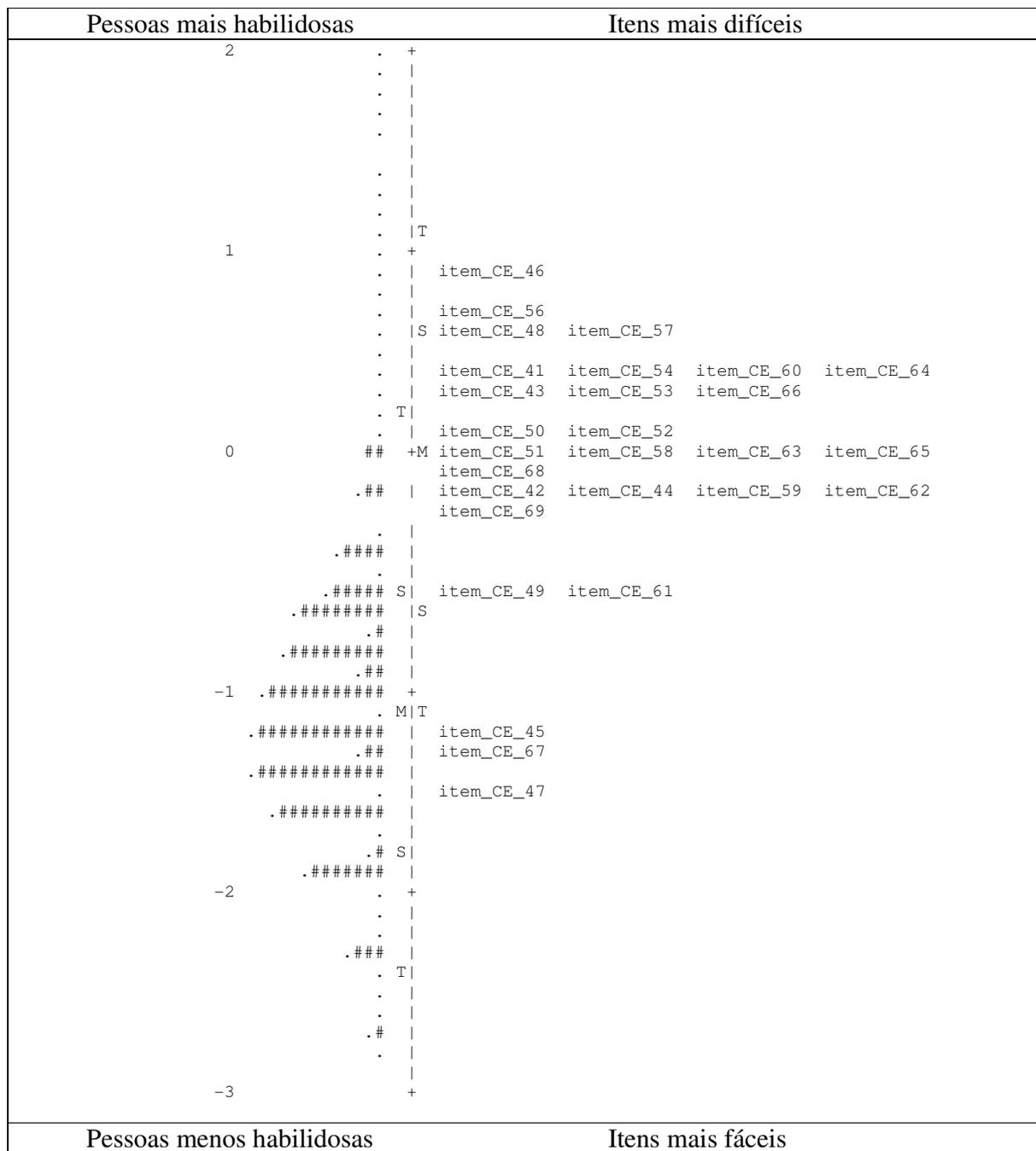


Figura 3. Comparação de habilidades dos estudantes e dificuldade dos itens.

As tabelas 12 e 13 contemplam as informações referentes aos parâmetros psicométricos dos itens da prova de Computação 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 0,91 de dificuldade (máximo) e -1,50 (mínimo). Também, no geral, os índices de ajuste de *Infit* e *Outfit* podem ser considerados suficientes. Quanto às pessoas, a maior parte evidenciou habilidades em pouco abaixo à da média dos itens (-1,12), o que pode indicar que a prova possui um nível de dificuldade maior que à habilidade das pessoas. Com relação às habilidades dos estudantes, houve pessoas com habilidades elevadas, a mais alta alcançou o nível de 1,92 de theta. Do mesmo modo, houve pessoas com baixa habilidade (theta = -3,47).

Tabela 12. Sumário das estatísticas fornecidas pelo modelo de Rasch por itens

Parâmetro dos itens				
Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	3.871,0	0,00	1,00	1,02
Desvio Padrão	1.562,4	0,56	0,05	0,08
Máximo	8.269	0,91	1,09	1,22
Mínimo	1.829,0	-1,50	0,91	0,87

Tabela 13. Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas
Parâmetro das pessoas

Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	7,1	-1,12	1,0	1,01
Desvio Padrão	3,1	0,64	0,13	0,21
Máximo	23,0	1,92	1,58	2,41
Mínimo	1,0	-3,47	0,34	0,29

A seleção dos itens sugeriu que os itens apresentassem boas propriedades psicométricas. Os itens do componente de Formação Geral que serviram de itens âncoras, ou seja, itens que fazem parte da prova 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 pelo colegiado dos cursos.

Escolha dos itens do Componente Específico

A escolha dos itens do Componente Específico que compuseram o PAC 2010 de Computação seguiu os mesmos requisitos utilizados na escolha dos itens de Formação Geral. Ressalta-se que foram indicados 10 itens para apreciação do colegiado do curso de Ciência da Computação com relação à forma e ao conteúdo da questão e, após análise do coordenador e professores do curso, os itens foram mencionados como pertinentes e suficientes para os propósitos da avaliação. Esse mesmo colegiado selecionou outras 20 questões para completar a prova, retiradas da prova do ENADE 2008 de Computação.

5.2 Estudo 2 – Análise da prova PAC de Computação 2010 equalizada ao ENADE 2005

5.2.1 Método

Fonte de dados

Este estudo contou com um banco de dados contendo informações referentes às respostas dos estudantes do curso de Ciência da Computação a uma Prova PAC equalizada ao ENADE 2005. Essas respostas contemplam dois componentes referentes a prova de Formação Geral e a prova de Conhecimento Específico, totalizando 40 questões.

Participantes

A amostra desta pesquisa foi composta por 60 alunos matriculados no curso de Ciência da Computação de uma universidade particular do interior do Estado de São Paulo, de ambos os sexos, em diferentes momentos do curso, distribuídos em 3 grupos a saber, ingressantes, intermediários e concluintes (Tabela 14). O critério para a distribuição dos alunos, descrita na Tabela 14, é semelhante aos critérios utilizados pelo MEC diferindo, apenas, no que concerne aos alunos considerados intermediários, já que essa faixa dos estudantes não é avaliada pelo ENADE.

Tabela 14. *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, alunos 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 15 mostra a distribuição de estudantes por diferenciação acadêmica, em que a maioria de 56,7% é considerada Concluente, de acordo com os critérios expostos na Tabela 15.

Tabela 15. *Distribuição dos estudantes pela situação acadêmica*

Situação Acadêmica	Número de Estudantes	%
Ingressantes	-	-
Intermediarios	26	43,3
Concluintes	34	56,7
Total	60	100

Material

Seruiu de material a prova do PAC de Ciência da Computação 2010, constituída com o formato semelhante ao ENADE, composta por 40 itens subdivididos em, 10 itens de Formação Geral, comum à todos os cursos, sendo 05 desses itens comuns ao ENADE de Computação 2005; 10 itens de Componente Específico comuns da prova do ENADE 2005; 20 itens de Componente Específico elaborados pelo colegiado de curso.

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.

5.2.2 Resultados e Discussão

O estudo 2 teve por objetivo a análise da prova PAC aplicada aos estudantes de Ciência da Computação, 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 metodologia semelhante a etapa 1, em que os dados foram tratados para que fosse possível sua utilização ao software estatístico Winsteps. Assim, a tabela 16 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 16. Estatísticas dos itens de Formação Geral no PAC de Ciência da Computação calibrados pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-total	Displace
01	19	21	-0,45A	0,75	0,62	0,34	-0,37
02	18	21	-0,07A	0,90	1,04	0,33	-0,20
03	20	21	-1,62A	1,14	0,96	0,14	-0,01
04	13	21	1,17A	0,82	1,07	0,59	0,22
05	18	21	-0,33A	1,09	1,37	0,28	0,07
06	15	21	0,82	0,93	0,86	0,50	0,00
07	11	21	1,91	0,88	0,79	0,60	0,01
08	06	21	3,38	1,01	0,95	0,56	0,01
09	15	20	0,59	1,39	1,32	0,24	0,00
10	12	21	1,65	0,90	0,81	0,57	0,01
Média	14,7	20,9	0,71	0,98	0,98		
DP	4,1	0,3	1,35	0,18	0,22		

Conforme Tabela 16, 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 estudante 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 8 ($b = 3,38$), retirada do componente de formação Geral do ENADE 2008, segundo os propósitos do colegiado do curso de Ciência da Computação.

A figura 4 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 4 percebe-se que as pessoas com níveis de habilidade de aproximadamente -1,1 respondem também as demais categorias, ou seja, respondem também as 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 0,1 em que a probabilidade é de 1,0 de acerto.

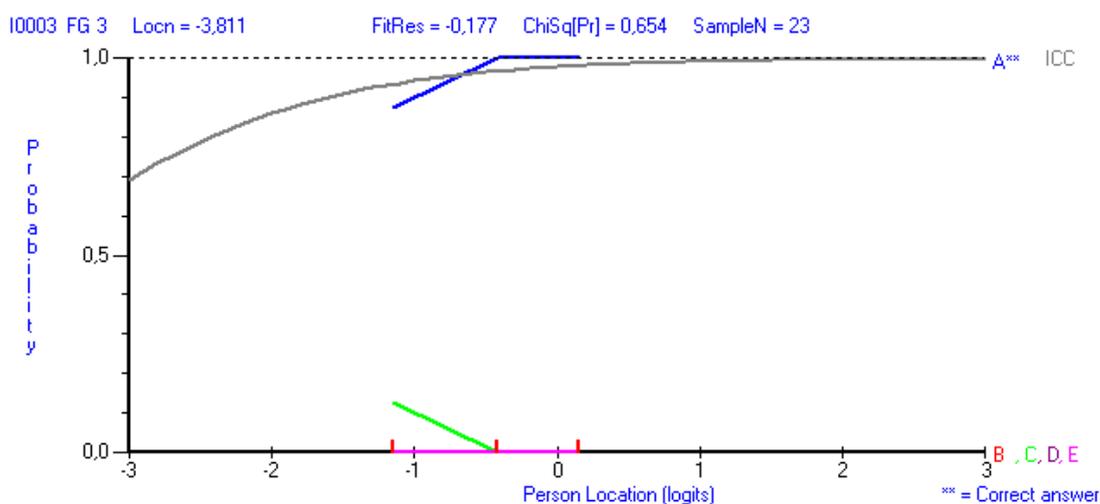


Figura 4. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 3 da prova de Formação Geral do PAC.

Já para a questão de número 8, a Figura 5 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 (A) é pouco provável em praticamente todos os níveis de habilidade dos sujeitos. 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, 0,65. A categoria de resposta C é a mais provável 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.

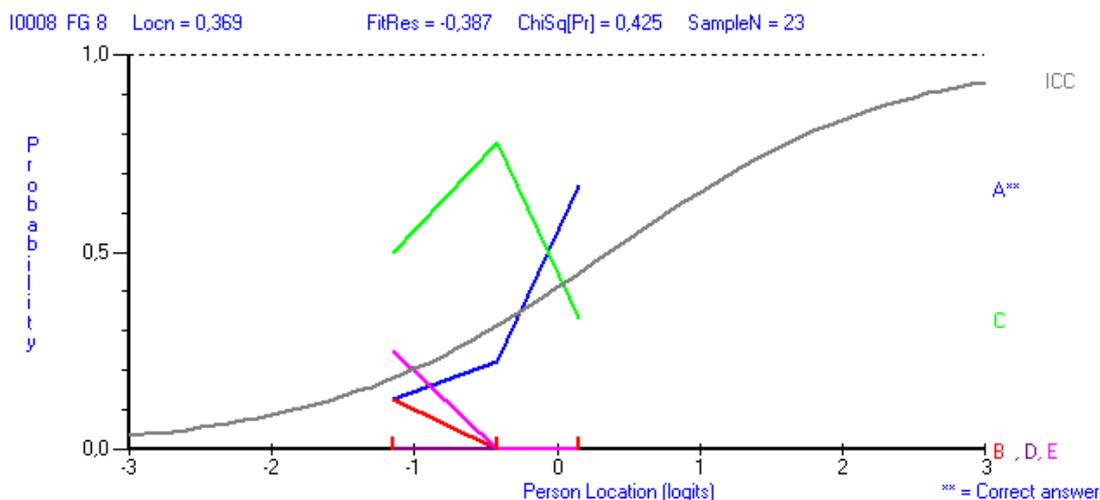


Figura 5. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 8 da prova de Formação Geral do PAC.

Com relação aos itens âncoras, indicado na tabela 16 pela letra “A”, estes apresentam medidas adequadas em que, a correlação item-total variou entre 0,28 a 0,59, a média *Infit* foi de 0,98 (DP= 0,18) e a de *Outfit* de 0,98 (DP= 0,22). 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 mais difícil apresentou índice de dificuldade de 1,17 (item 4).

Os índices de ajuste ao modelo de Rasch referentes aos itens do Componente Especifico do PAC Ciência da Computação estão dispostos na Tabela 17. Considerando a prova como um todo, isto é, não distinguindo entre itens ancoras e itens comuns, houve

itens que podem ser considerados inadequados no que se refere às suas medidas. A medida de correção item-total, por exemplo, obteve valores abaixo do mínimo esperado que é de 0,20 nas questões 16 ($r_{item=total}=0,09$), 30 ($r_{item=total}=0,02$), 35 ($r_{item=total}=0,04$) e 39 ($r_{item=total}=0,19$). Na medida de *Outfit* também se manifestaram valores que indicam padrões de resposta inesperado como a questão 35 (*Outfit*= 1,46), acima do valor crítico de 1,40.

Tabela 17. Estatísticas dos itens no Componente Específico do PAC de Ciência da Computação calibrados pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	<i>Infit</i>	<i>Outfit</i>	Correlação Item-toal	Displace
11	8	21	-0,34	1,04	0,99	0,27	0,00
12	9	21	-0,56	0,81	0,78	0,58	0,00
13	3	21	1,06	1,02	0,87	0,22	0,00
15	10	21	-0,49A	0,91	0,86	0,52	-0,28
16	3	21	1,06	1,10	1,03	0,09	0,00
17	7	21	0,02A	1,00	1,15	0,31	-0,13
19	7	21	-0,52A	0,99	0,95	0,26	0,40
21	15	21	-1,89	1,03	0,97	0,28	-0,01
23	8	21	-0,34	1,04	1,00	0,27	0,00
24	5	21	0,40	1,00	0,90	0,29	0,00
25	13	21	-1,41	1,02	1,00	0,30	-0,01
27	5	21	0,40	1,00	1,26	0,20	0,00
28	10	21	-0,77	0,90	0,87	0,46	0,00
29	6	21	0,13	1,02	1,05	0,23	0,00
30	5	21	0,40	1,13	1,40	0,02	0,00
31	15	21	-1,89	0,91	0,97	0,40	-0,01
32	5	21	0,40	1,01	1,10	0,22	0,00
35	5	21	0,40	1,09	1,46	0,04	0,00
36	4	21	0,70	0,96	0,83	0,33	0,00
38	5	21	0,40	1,00	0,89	0,29	0,00
39	8	21	-0,34	1,09	1,06	0,19	0,00
40	4	21	0,70	0,95	0,86	0,32	0,00
Média	7,3	21	-0,11	1,00	1,01		
DP	3,5	0,0	0,82	0,07	0,17		

A prova do Componente Específico do PAC de Ciência da Computação apresentou itens que, de acordo com seus respectivos índices de dificuldade, podem ser considerados muito fáceis, fáceis, médios, difíceis e muito difíceis. Na Figura 6 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 13 ($b=1,06$) e 16 ($b=1,06$); Já a questão mais fácil da prova foram as de número 21 ($b= -1,89$) e 31 ($b= -1,89$).

Outra característica que pode ser observada na figura 6, 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 apresentou uma dificuldade um pouco acima da média de habilidade dos respondentes.

A curva característica referente ao item mais difícil da prova é apresentada na Figura 7, em que a probabilidade de acerto ao item corresponde ao eixo Y e a habilidade do sujeito 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 -0,5, mas tem uma queda e se estabiliza em aproximadamente 0,1. As respostas mais prováveis a partir de -0,5 de habilidade são dadas pela categoria D que cresce com o aumento da habilidade dos sujeitos, e a categoria C que também após -0,5 de habilidade começa a decrescer. Este é um item que requer especial atenção, uma vez que pode sugerir problemas na construção do mesmo.

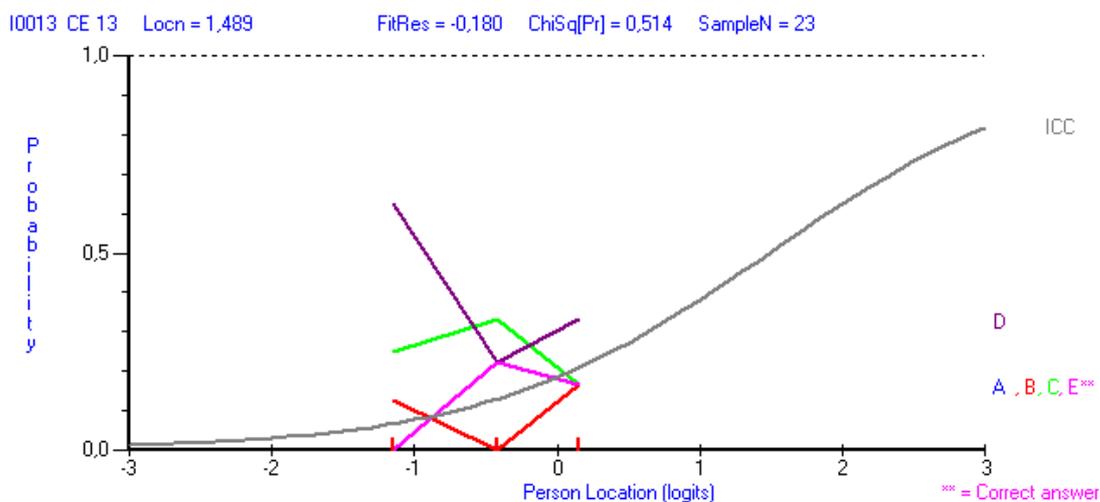


Figura 7. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 13 da prova de Componente Específico do PAC.

Para o item mais fácil da prova do Componente Específico da prova PAC, obteve-se a curva característica do item evidenciada na Figura 8. Assim, no item indicado como representante das competências mínimas a serem desenvolvidas pelo curso, percebe-se que pessoas com níveis de habilidade de aproximadamente -1,10 ainda tem quase 40% de

chance de responderem errado, com possibilidade de resposta na categoria B. Conforme a habilidade das pessoas aumenta, a resposta correta (A) passa a ser a mais provável para um nível de habilidade de até 0,10 em que a probabilidade é quase 1,00 de acerto.

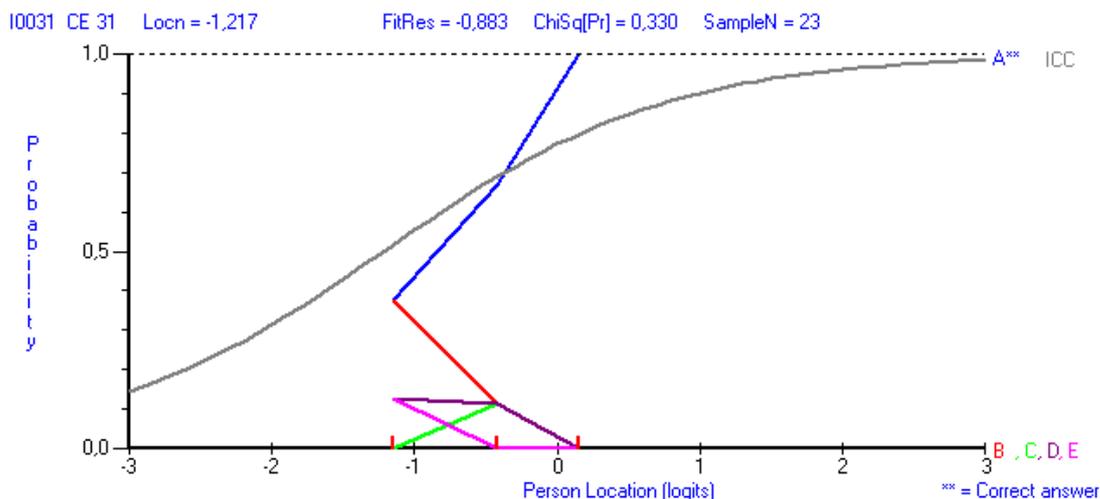


Figura 8. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 31 da prova de Componente Específico do PAC.

A Tabela 18 apresenta os parâmetros psicométricos dos itens das provas administradas no PAC de Ciência da Computação e a Tabela 19 os parâmetros psicométricos dos 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 (-0,11) e houve indicadores que alcançaram até 1,07 de dificuldade (máximo) e -2,20 (mínimo). Os índices de ajuste de *Infit* e *Outfit* estiveram bons, sendo que apenas um índice de *Outfit* foi igual a 1,47, que é o limite indicado de desajuste (1,40). Quanto às pessoas, a maior parte evidenciou habilidades um pouco abaixo da média dos itens (-0,88), o que indica, como já citado, que a prova possui um nível de dificuldade um pouco maior que a habilidade das pessoas. Não houve pessoas com habilidades muito

elevadas, e a mais alta alcançou o nível 0,12 de tetha. Em contrapartida, houve pessoas com baixa habilidade ($\theta = -2,20$).

Tabela 18. Sumário das estatísticas fornecidas pelo modelo de Rasch dos itens.

Parâmetro dos itens

Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	7,3	-0,11	1,00	1,01
Desvio Padrão	3,5	0,82	0,07	0,17
Máximo	15,0	1,07	1,13	1,47
Mínimo	3,0	-1,89	0,81	0,78

Tabela 19. Sumário das estatísticas fornecidas pelo modelo de Rasch por pessoas.

Parâmetro das pessoas

Estatísticas	Escore Bruto	Dificuldade	<i>Infit</i>	<i>Outfit</i>
Média	7,6	-0,88	1,00	1,01
Desvio Padrão	2,8	0,70	0,17	0,26
Máximo	12,0	0,12	1,37	1,81
Mínimo	3,0	-2.20	0,68	0,64

5.2.3 – Análise e interpretação dos resultados do PAC comparados ao ENADE em uma mesma escala

5.2.3.1 Resultados e discussão

A presente etapa teve por objetivo apresentar os resultados das análises psicométricas do componente específico da prova aplicada ao curso de Ciência da Computação do PAC e compara-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 curso.

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 20 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 (58,9%) quanto os concluintes (60,6%), foram categorizados no conceito 3 do ENADE, totalizando 59,7% de toda a amostra. Cabe ressaltar que os estudantes considerados Ingressantes cursavam, na ocasião do Exame, os 1º e 2º semestres do curso e, os considerados Concluintes, os 7º e 8º semestres da graduação (INEP, 2005).

Tabela 20. Distribuição de estudantes de Ciência da Computação por conceito no ENADE 2005

Conceito	Ingressante		Concluente		Total	
	N	%	N	%	N	%
1	95	01,3	65	01,0	160	01,1
2	589	08,2	600	08,9	1189	08,5
3	4254	58,9	4078	60,6	8332	59,7
4	1374	19,0	1252	18,6	2626	18,8
5	383	05,3	397	05,9	780	05,6
SC	526	07,3	342	05,1	868	06,2
Total	7221	100,0	6734	100,0	13955	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. Para a obtenção de uma escala comum entre os estudantes do ENADE de Ciência da Computação do ENADE 2005 e os estudantes do PAC de Ciência da Computação, 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 e 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 do PAC com o ENADE é possível ter como grupo de referência todos os estudantes do Brasil que fizeram o ENADE no ano de 2005 – as notas percentílicas tem como grupo de referencia a amostra brasileira do ENADE e, com essa equalização, os valores referentes à escala PAC correspondem ao

mesmo nível de habilidade no ENADE. A distribuição dos percentis de nota média no ENADE, divididos por ingressantes e concluintes, é apresentada na Figura 9.

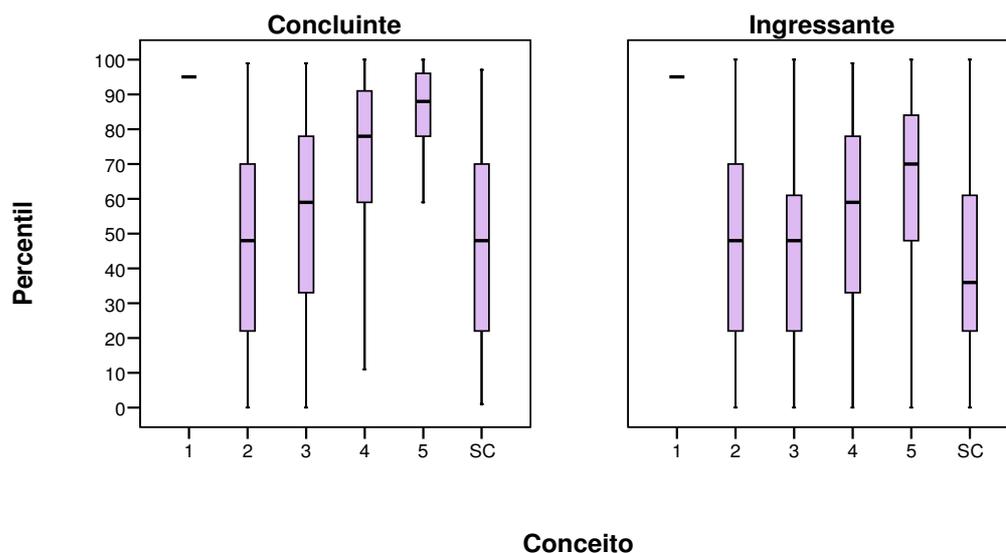


Figura 9. Box-plot dos percentis de nota média no ENADE de ingressantes e concluintes por conceito do curso.

A análise da Figura 9 permite dizer que, de acordo com os percentis obtidos, o desempenho dos estudantes que participaram do ENADE de Ciência da Computação 2005 aumenta conforme aumenta o conceito dos cursos. Desta forma, tanto 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.

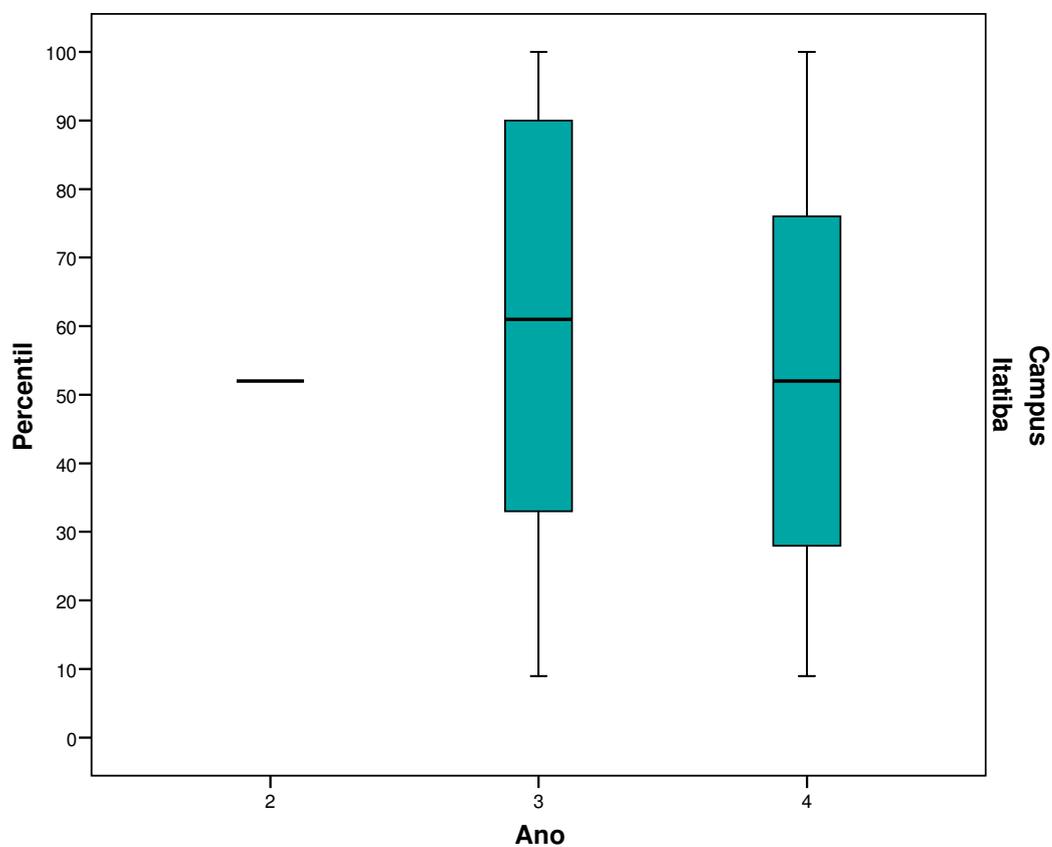


Figura 10. Box-plot dos percentis de nota média no PAC por ano.

A Figura 10 contempla os percentis de nota média dos estudantes do PAC, divididos por ano. 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 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 10, mediana em torno de percentil 60 e a nota máxima alude ao percentil 99. Essas

mesmas medidas para os estudantes do 4º ano denotam nota mínima congruente com o percentil 10, mediana em torno de 52 e a nota máxima por volta do percentil 99.

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) do curso, devido a não possuir alunos matriculados. Já com relação ao desempenho dos alunos concluintes do PAC, matriculados no 4º ano da graduação, em comparação aos alunos concluintes do ENADE, os primeiros obtiveram notas médias percentílicas semelhantes ao desempenho dos alunos com Conceito 3 no ENADE.

Os processos de equalização de provas de habilidades acadêmicas que buscam verificar o desempenho de estudantes ainda são poucas. Na literatura especializada, estão citados esses processos, com dados advindos do SAEB (Sistema Nacional de Avaliação da Educação Básica). Nesse sentido, os dados do SAEB, desde a década de 1990, vem se utilizando da TRI como metodologia de equalização e análise de dados. Assim, calculando os parâmetros de dificuldade, discriminação e o acerto ao acaso os itens considerados com boas propriedades psicométricas são armazenadas em um Banco Nacional de Itens e podem ser utilizados na situação de avaliações nacionais e estaduais (Locatelli, 2002). A autora ressalta as vantagens da equalização quando se pretende colocar os resultados dos testes em uma mesma escala, possibilitando a comparação entre diferentes grupos de estudantes. Informações mais detalhadas sobre equalização de provas com os dados do SAEB estão disponíveis nos estudos de Klein (2009) e Souza (2008).

Na avaliação do Ensino Superior que pratica 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 *Stricto Sensu* da Universidade São Francisco. Dentre as pesquisas realizadas pelo LabAPE está a realização de provas

institucionais construídas de modo semelhante a metodologia apresentada nesta pesquisa desde a criação 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 desempenho 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).

Estudo 3 – Análise do mapa de itens

A elaboração de um mapa de itens para uma prova que avalia habilidades de estudantes permite a demonstração de itens com comprovada dificuldade ou resultados que sugerem erros na construção daquele item. A análise foi realizada para a prova PAC 2010 de Computação, descrita a seguir.

Sendo a prova PAC 2010 de Computação composta por 40 questões, distribuídas em 10 de Formação Geral e 30 de Componente Específico, a análise do mapa de itens contemplará apenas as questões referente ao Componente Específico. Para facilitar a leitura, foram selecionados quatro itens considerados difíceis e quatro com grau de dificuldade considerado fácil.

Conforme apresentado no capítulo anterior, os itens considerados mais difíceis para o componente específico foram questões 13 ($b=1,06$), 16 ($b=1,06$), 36 ($b=0,70$) e 40 ($b=0,70$). Abaixo será tratado o que o item estava propondo e qual fator elevou o grau de dificuldade.

A Figura 11 representa a curva característica do item 13 e demonstra que a questão apresentou problemas na sua construção, pois a resposta correta (E) possui uma crescente

até o nível de habilidade $-0,5$ e não mantem a crescente de acordo com o nível de habilidade. O item exigia a habilidade de possuir uma visão sistêmica e integral da área de computação, e versava sobre a temática de alocação de memória de dados, no sistema interno de um computador. As opções de escolha da alternativa correta eram baseadas na interpretação de afirmações e relaciona-las com o conteúdo apresentado na contextualização da pergunta. Por se tratar de uma temática naturalmente complexa, o relacionamento com as afirmações provavelmente foram os responsáveis pelo grau elevado de dificuldade do item.

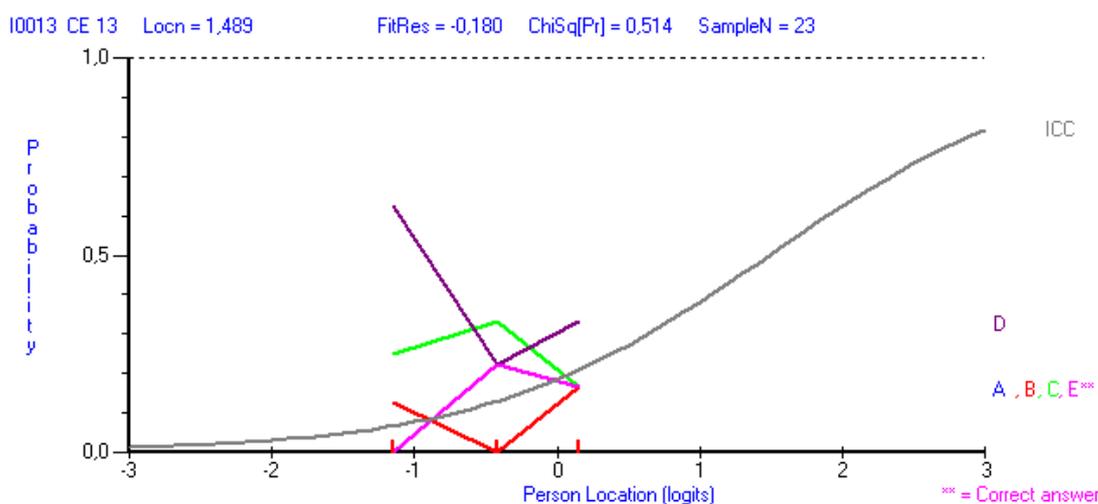


Figura 11. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 13 da prova de Componente Específico.

A Figura 12 representa a curva característica do item 16 e demonstra que a questão apresentou a categoria E como mais provável de resposta, mas a categoria correta é a C, este item também sugere uma reflexão quanto a sua construção, uma vez que a curva indica que pessoas mais habilidosas optam pela categoria errada. O item exigia a habilidade de

saber modelar e especificar soluções computacionais para diversos tipos de problemas, e tratava da temática principal da área de computação, que diz respeito à lógica. As opções de respostas foram apresentadas por meio de uma imagem e a resposta correta segue uma linha de raciocínio contrária a opção de maior probabilidade.

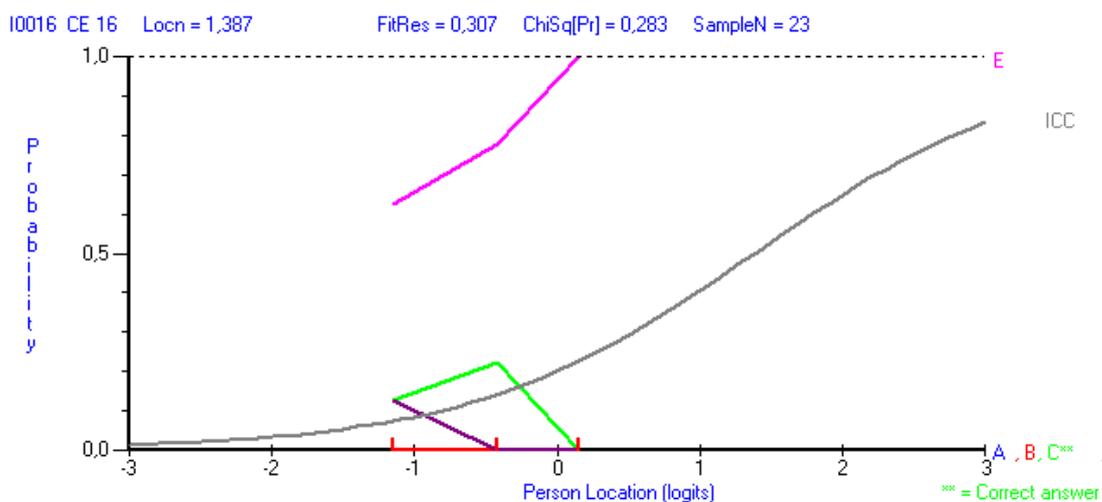


Figura 12. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 16 da prova de Componente Específico.

A Figura 13 representa a curva característica do item 36 e demonstra que a questão representou um grau de dificuldade maior que a habilidade dos sujeitos. Sua construção foi bem elaborada, mas para responde-la era necessário associação das afirmações, novamente os sujeitos demonstraram certa dificuldade para responder corretamente. O assunto abordado na questão trata de mecanismos de busca dentro de um sistema de informação, onde as afirmações que subsidiaram as categorias de respostas, apresentaram variações pequenas, o que deve ter contribuído para o aumento da dificuldade.

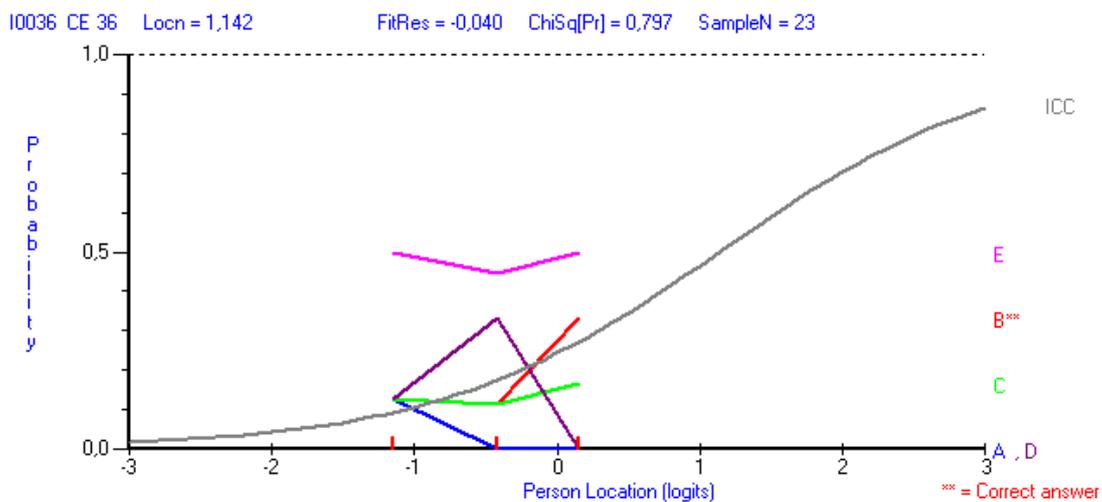


Figura 13. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 36 da prova de Componente Específico.

A Figura 14 representa a curva característica do item 40 e demonstra que a questão apresentou problemas na sua construção, pois a resposta correta (E) não representou a escolha das pessoas de maior nível de habilidade. O item foi construído pela apresentação de uma figura e suas opções de resposta eram apresentadas em expressões matemáticas. Para responder a categoria correta era necessário o conhecimento prévio de expressões matemáticas e a simbologia computacional. O tema referente ao item trazia conceitos de arquiteturas de redes de computadores.

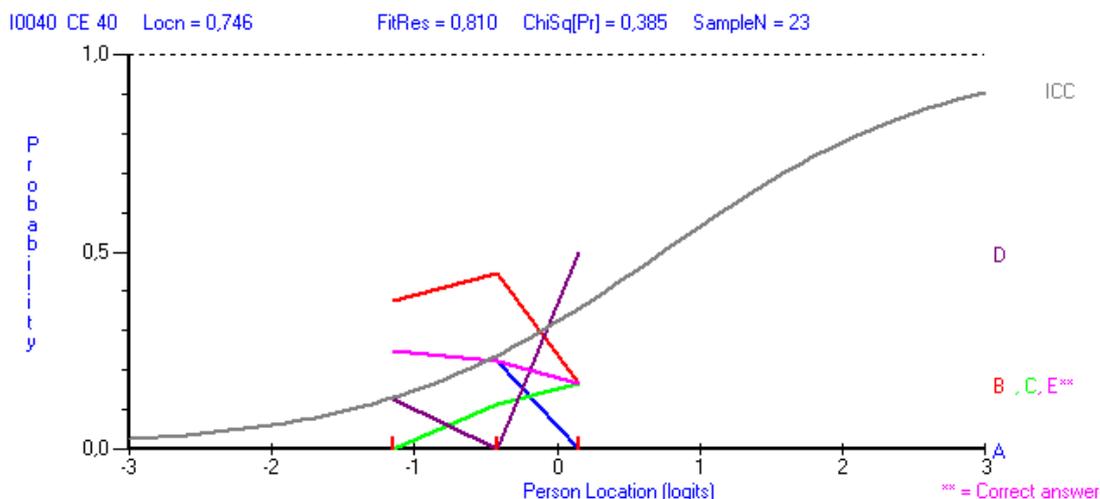


Figura 14. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 40 da prova de Componente Específico.

Na sequência serão apresentados os quatro itens do componente específico considerados mais fáceis, a saber, 28 ($b=-0,77$), 25 ($b=-1,41$), 21 ($b=-1,89$) e 31 ($b=-1,89$). Abaixo será tratado o que o item estava propondo e qual fator elevou o caracterizou como um item fácil.

A Figura 15 representa a curva característica do item 28 e já é possível constar que pessoas com nível de habilidade menor que -1 tem a possibilidade de pouco mais de de 10% de acerto e ainda próximo desta faixa de habilidade os sujeitos escolhem as categorias A e B, mas quando o nível de habilidade atinge o nível 0 a categoria escolhida passa a ser a correta. Este item também pode ser considerado como bem elabora e as possíveis justificativas para que o mesmo seja caracterizado como fácil é o fato de ser enquadrado como teste de mesa, onde a expressão faz parte do item e o teste é expressamente demonstrado, isso também permite a conclusão que exista a confusão pela escolha das categorias A e B pelo simples fato de um erro matemático.

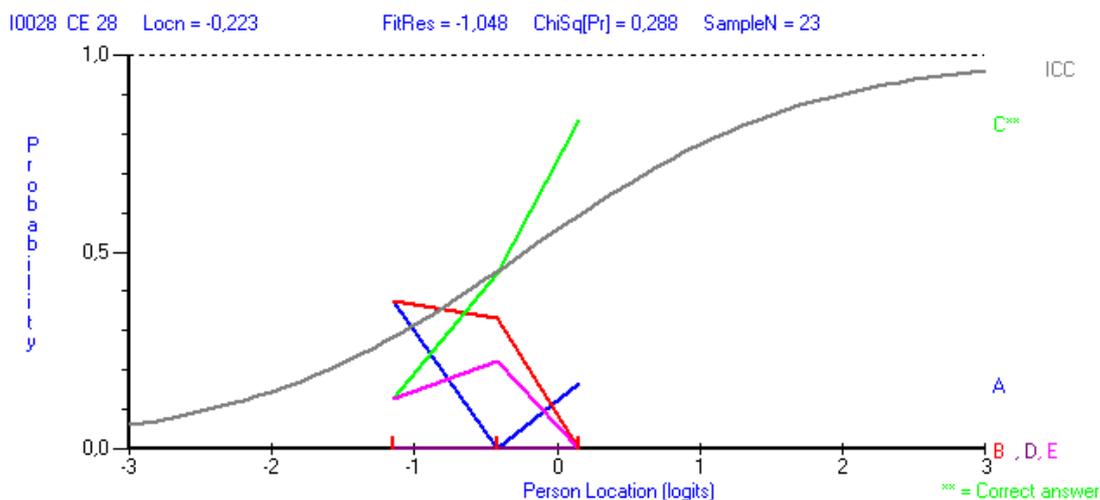


Figura 15. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 28 da prova de Componente Específico.

A Figura 16 representa a curva característica do item 25 e novamente é possível observar que pessoas com nível de habilidade menor que -1 tem a possibilidade de acerto de praticamente 50%. Este item foi bem elaborado, sua contextualização poderia elevar o grau de dificuldade, mas o assunto abordado no item é atual e teoricamente faz parte do dia a dia dos estudantes de computação, acrescido de categorias de opção de respostas óbvias, contribuíram para que esse item se torna-se fácil.

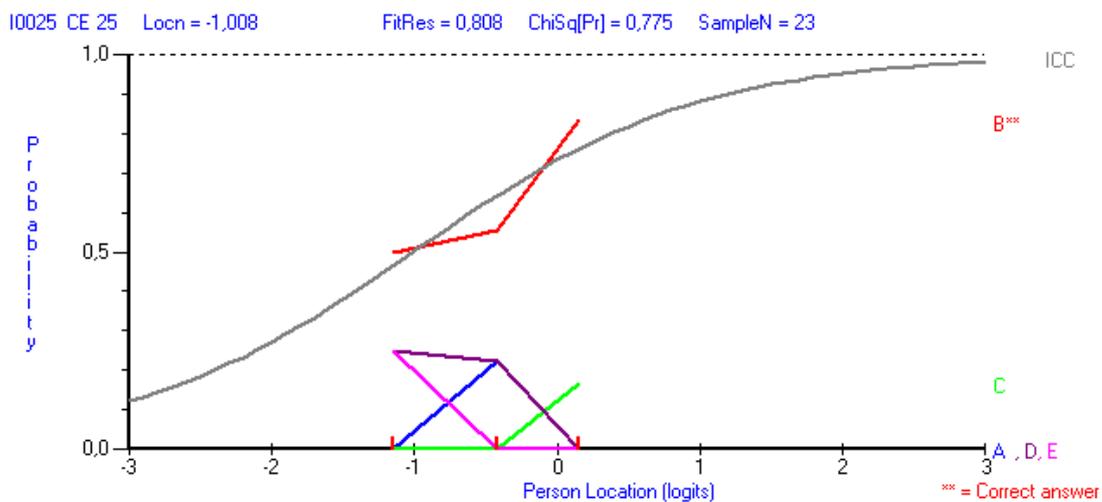


Figura 16. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 25 da prova de Componente Específico.

A Figura 17 representa a curva característica do item 21 e demonstrou que mesmo a questão sendo considerada fácil, quando a habilidade do sujeito está próxima a 0,1, apresentou-se uma queda da escolha pela categoria correta C pela categoria B. Podemos concluir que mesmo se tratando de uma temática simples, ou seja, tecnologias de armazenamento, para seleção das categorias era necessário estipular um relacionamento entre quatro afirmações o que pode ter confundido as pessoas de habilidade um pouco mais elevada.

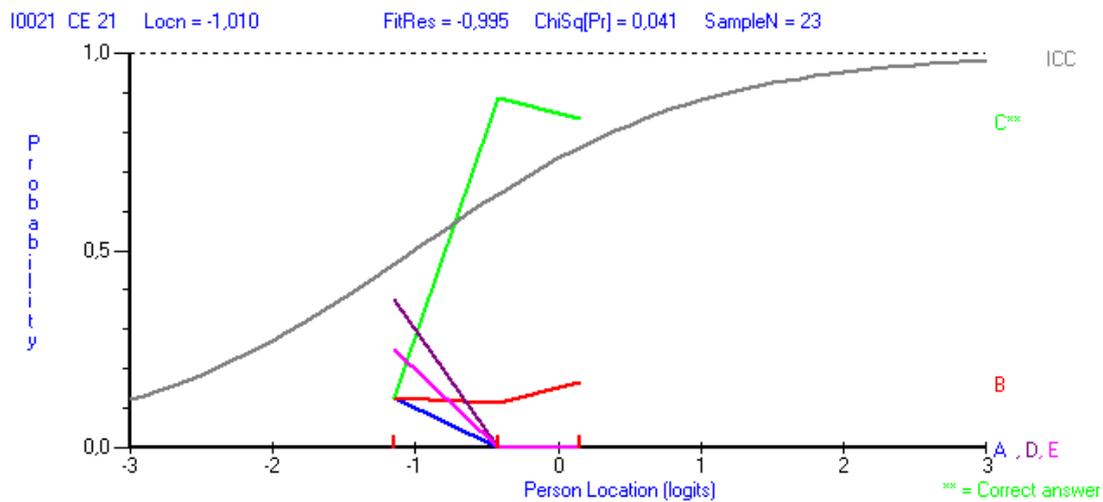


Figura 17. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 21 da prova de Componente Específico.

A Figura 18 representa a curva característica do item 31 e já é possível constatar que pessoas com nível de habilidade menor que -1 tem a possibilidade praticamente 30% de acerto e quando o nível de habilidade atinge o nível 0 esta passa a ser a categoria escolhida. O item possui duas características que o sugerem como fácil, primeiro por que se trata de uma questão de resposta direta, não necessitando muito raciocínio do sujeito no momento da escolha da resposta e segunda por se tratar de um assunto básico da área de computação, que pode ser considerado como linguagem natural não muito técnico.

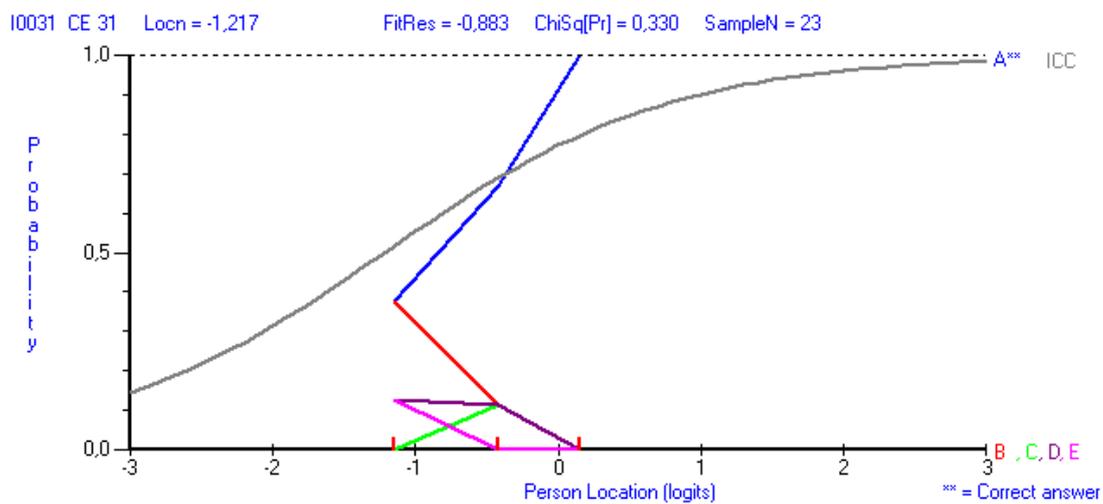


Figura 18. Relação entre habilidade das pessoas e probabilidade de resposta por categoria no item 31 da prova de Componente Específico.

CONCLUSÃO E CONSIDERAÇÕES FINAIS

O propósito principal do presente estudo somente foi possível, graças à expertise adquirida e transmitida pelo LabAPE (Laboratório de Avaliação Psicológica e Educacional). A equalização da prova de Ciência da Computação do ENADE 2005 com a prova PAC apropriou-se das propriedades matemáticas propostas pela Teoria de Resposta ao Item. Ainda foram abordadas as questões de habilidades acadêmicas, sua origem e trajeto, seus principais estudiosos, além de retomadas as questões acerca do modelo CHC das habilidades cognitivas, tendo em vista a premissa colocada por McGrew (2008) de que o modelo não deve ser visto como um fim, mas que sugere novas investigações psicométricas sobre as estruturas das capacidades humanas.

O acréscimo de mais seis habilidades cognitivas ao modelo CHC permitiu a ampliação e o aprimoramento das maneiras de conceber a inteligência em adultos, possibilitando novas pesquisas. Das habilidades acrescidas ao modelo, a mais importante para este estudo é a Gkn, haja vista sua importância no desenvolvimento de competências relacionadas ao bom desempenho em tarefas acadêmicas. Desse modo, o domínio do conhecimento especializado, uma das características do Gkn, comprova o sucesso em provas de habilidades acadêmicas como, por exemplo o ENADE.

Questões históricas da avaliação da Educação Superior no Brasil também foram abrangidas no presente estudo, 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 como 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 ligados a outras áreas do conhecimento. Esse objetivo está presente desde a instituição do SINAES e no Manual do ENADE do ano de 2005, disponível para consulta eletrônica no sítio do INEP (www.inep.gov.br).

Em sua trajetória de nove anos, o ENADE recebeu diversas críticas, algumas referenciadas no presente estudo, identificando que em suas edições anteriores alguns de seus propósitos foram perdidos. Em algumas situações foram motivadas por propagandas marqueteiras de Instituições de Ensino Superior que utilizaram os conceitos para captação de novos alunos ou por mudanças no cenário político-econômico. Parte dessa problemática também foi tratada no presente trabalho.

Voltando ao foco principal do presente estudo, pode-se constatar que o modelo matemático proposto pela TRI mostrou-se bem sucedido para 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 habilidades dos itens quanto para a escolha de itens âncoras que compuseram a prova de Ciência da Computação do Programa de Avaliação Continuada (PAC).

Ainda no mesmo estudo, pode-se identificar a dificuldade da prova PAC de Ciência da Computação, os itens mais fáceis e mais difíceis e, ainda, o desempenho dos estudantes que realizaram essa prova. Por fim, os resultados do Estudo 2, isto é, o resultado da equalização da prova, permitiram a comparação entre os estudantes que fizeram a prova

PAC de Ciência da Computação e aqueles que fizeram a prova de Ciência da Computação 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.

Nos resultados apresentados pela análise do mapa de itens, conclui-se que estudos similares permitem que os resultados indiquem se o item demonstra o real nível de dificuldade ou se apresentou algum problema em sua construção. Nesse contexto foi possível identificar que alguns níveis de dificuldade representavam os níveis esperados; em contrapartida, alguns itens apresentaram problemas em sua construção.

Ao final da pesquisa, em toda temática que envolve a questão da Avaliação da Educação Superior e a Teoria de Reposta ao Item, buscou-se construir valores e corroborar as teorias existentes, não desmerecendo a TCT, mas sim, ressaltando as vantagens do modelo matemático da TRI. Nesse sentido, os instrumentos de avaliação pautados nas respostas psicométricas têm muito a contribuir na elaboração de instrumentos eficazes que tragam resultados positivos para o sistema de avaliação da educação de nosso país.

REFERÊNCIAS

Ackerman, P. L., & Lohman, D. F. (2006). Individual differences in cognitive function. Em Patricia A. Alexander & Philip H. Winne (Orgs.). *Handbook of educational psychology*, (pp. 139-160). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: knowledge, abilities, and nonability traits. *Psychology and Aging*, 14(2), 314-330.

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, C. R. (2000). *Teoria de Resposta ao Item: conceitos e aplicações*. São Paulo: Sinape.

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*. Recuperado em abril de 2011. Disponível em <http://www6.senado.gov.br/legislacao/DetalhaDocumento.action?id=142301>.

Brasil (1996). Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. *Diário Oficial da República Federativa do Brasil*. Disponível em http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm. Recuperado em abril de 2011.

Brasil. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – MEC/INEP (2007). *SINAES – Sistema Nacional de Avaliação da Educação Superior: da concepção à regulamentação* (4ª. ed., ampl). Brasília: MEC/INEP.

Brito, M. R. F. (2008). O SINAES e o ENADE: da concepção à implantação. *Revista Avaliação, 13*(3), 841-850.

Brito, M. R. F., 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*. Oakleigh: Cambridge University Press.

Castro, M. H. G. (2000). Sistemas nacionais de avaliação e de informações educacionais. *São Paulo em perspectiva, 14*(1), 121-128.

Castro, M. H. G. (2009). A Consolidação da Política de Avaliação da Educação Básica no Brasil. *Meta: Avaliação, 2*(3), 271-296.

Castro, E. B. & Sá, M. A. D. (2002). Habilidades, Competências, Valores e Atitudes: Um perfil para o profissional de computação e informática. Em XXII Congresso da Sociedade Brasileira de Computação, 2002, Florianópolis/SC. Anais do XXII Congresso da SBC. Florianópolis: Sociedade Brasileira de Computação, 2002. v. 4. p. 115-126.

Conde, F. N., & Laros, A. J. (2007). Unidimensionalidade e a propriedade de invariância das estimativas da habilidade pela TRI. *Avaliação Psicológica 6*(2), 205-215.

- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de dados*. Porto Alegre: Bookman.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications.
- Hambrick, D. Z. (2005). The role of domain knowledge in higher-level cognition. Em O. Wilhelm & R.W. Engle. *Handbook of understanding and measuring intelligence* (pp.361-372). Sage Publications, Inc.
- Klein, R. (2009). Utilizando da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB). *Meta*, 1(2), 125-140.
- McGrew, K. S. (2004). *Cattell-Horn-Carroll (CHC) definition Project*. Disponível em <http://www.iapsych.com>. Recuperado em abril de 2011.
- 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.
- 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 factor analytic research*. Disponível em: <http://www.iapsych.com>. Recuperado em abril de 2011.
- Muñiz, J. (2010). Las teorías de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo*, 31(1), 57-66.
- Nogueira, S. O. (2008). ENADE: Análise de itens de formação geral e de estatística pela TRI. Dissertação de Mestrado, Programa de Pós-Graduação *Stricto 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.
- 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.
- 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.
- 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., Vendramini, C. M., Taxa, F., Muller, F. A., Lukjanenko, M. F., & Sampaio, I. S. (2001). Competências e Habilidades Cognitivas: diferentes definições dos mesmos construtos. *Psicologia: Teoria e Pesquisa*, 17(2), 151-159.
- 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., 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., 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.

- Ribeiro, I. (1998). *Mudanças no desempenho e na estrutura das aptidões: Contributos para o estudo da diferenciação cognitiva em jovens*. Braga: Universidade do Minho.
- Schelini, P. W. (2006). Teoria das inteligências fluída e cristalizada: início e evolução. *Estudos de Psicologia*, 11(3), 323-332.
- 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.
- 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.
- Souza, S. Z. L., & Oliveira, R. P. (2003). Políticas de avaliação da educação e quase mercado no Brasil. *Educação e Sociedade* [online], 24(84), 873-895. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-73302003000300007&lng=en&nrm=iso>. Recuperado em: 23 outubro 2011.
- Tejada, A. J. R., & Meléndez, C. P. (2001). *Nuevos modelos para la medición de Actitudes*. Valencia: Promolibro.
- Valle, R. C. (2000). Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, 21, 7-91.
- Vendramini, C. M. M. (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. (2005). Aplicação da Teoria de Resposta ao Item na avaliação educacional. Em R. Primi (Org.). *Temas em Avaliação Psicológica* (pp. 229-254). São Paulo: Casa do Psicólogo.

Vendramini, C. M. M. (2006). Análise Multidimensional do Exame Nacional de Desempenho do Estudante. Em D. Ristoff e A. Limana, M. R. F. Brito. (Org.). Enade: Perspectiva de avaliação dinâmica e análise de mudanças (1ª ed.). Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, v. 2, p. 45-58.

Verhine, R. E., & Dantas, L. M. V. (2005). *Avaliação da Educação Superior no Brasil: do Provão ao ENADE*. Documento Preparado para o Banco Mundial. Disponível em: <<http://www.isp.ufba.br/avalia%C3%A7%C3%A3o%20da%20Ed%20Superior%20do%20Provao%20ao%20ENADE.pdf>>. Acesso em: 15 outubro 2011.

Vianna, H. M. (2003). Avaliações Nacionais em Larga Escala: análises e propostas. *Estudos em Avaliação Educacional*, 27, (jan-jun), 41-76.