

MARJORIE CRISTINA ROCHA DA SILVA



VALIDADE DO ENADE PARA AVALIAÇÃO DA QUALIDADE
DOS CURSOS: MODELO DE RASCH E MULTINÍVEL

ITATIBA

2011

MARJORIE CRISTINA ROCHA DA SILVA

VALIDADE DO ENADE PARA AVALIAÇÃO DA QUALIDADE
DOS CURSOS: MODELO DE RASCH E MULTINÍVEL

Tese apresentada ao Programa de Pós-Graduação
Stricto Sensu em Psicologia da Universidade São
Francisco para obtenção de título de Doutor em
Psicologia.

ORIENTADOR: PROF. DR. RICARDO PRIMI

ITATIBA

2011

UNIVERSIDADE SÃO FRANCISCO

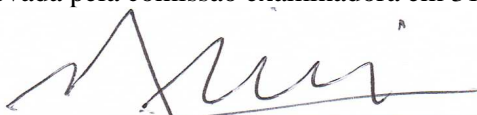
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM PSICOLOGIA

DOUTORADO

VALIDADE DO ENADE PARA AVALIAÇÃO DA QUALIDADE

DOS CURSOS: MODELO DE RASCH E MULTINÍVEL

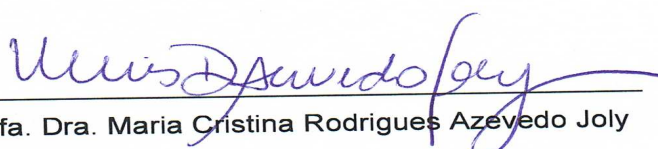
Este exemplar corresponde à redação final da tese de doutorado defendida por Marjorie Cristina Rocha da Silva aprovada pela comissão examinadora em 31/03/2011.



Prof. Dr. Ricardo Primi
Presidente



Profa. Dra. Claudette Maria Medeiros Vendramini



Profa. Dra. Maria Cristina Rodrigues Azevedo Joly



Prof. Dr. Robert Evan Verhine



Prof. Dr. Carlos Henrique Sancineto da Silva Nunes

ITATIBA

2011

Assim adquires a sabedoria que te dá a vida.

Se a adquirires, terás um bom futuro
e a tua esperança não será frustrada.

Provérbios 24:14

Agradecimentos

Foram muitas as dificuldades encontradas pelo caminho para se chegar até aqui, mas nenhuma delas é equiparável as conquistas, em termos de conhecimento, experiência profissional e principalmente, as grandes amizades adquiridas e aperfeiçoadas nesse percurso. Gostaria primeiramente de agradecer a Deus pelo modo como tem me sustentado e cuidado de mim em todo o tempo, sem Ele nada disso teria sentido.

Agradeço imensamente aos meus pais, Maria Teresa e Hélio, que sempre me apoiaram a persistir em meus sonhos com garra, maturidade, sabedoria e principalmente, verdade e caráter.

Ao meu esposo José Rocha, que se permitiu muitas vezes estar em segundo plano para que eu pudesse terminar meus trabalhos, mesmo durante as madrugadas. E que agüentou com paciência as dificuldades de cada momento, sem por isso diminuir seu amor e cuidado por mim.

Ao meu filho Nicolas, que com seu jeito meigo e ingênuo me dava parâmetros do quanto eu podia me dedicar a outras coisas que não a ele e a família, com frases do tipo “mãe, sei que você precisa trabalhar mas sinto muitas saudades”.

Ao meu filho Isaque, que acompanhou esse último ano de doutorado “dentro de mim” e me presenteou estando em meus braços antes da concretização dessa jornada! Obrigado filho pelo presente especial de seu nascimento em meio a tanta correria e, por, juntamente com o Nicolas, me proporcionar visualizar ainda mais meus sonhos para o futuro.

Aos amigos e parentes, ‘extra-acadêmicos’, que apesar do maior distanciamento nesse período, não deixaram de me apoiar. Especialmente gostaria de citar alguns nomes e suas respectivas famílias: Gilmara, Rebeca, Rose, Shirlei, Ruth, Dinah, Marlice, Anita e

Rogério, André, Vânia, Tia Sônia, Sanny Padovany, Mariana, Pr. Edmar e Fabiana, Cris e Rodrigo, Marta, entre tantos outros tão especiais quanto eles, mas que nesse espaço seria difícil citar todos.

Gostaria de agradecer especialmente ao meu orientador, Ricardo Primi, pelos conhecimentos, amizade e mesmo críticas, que me fizeram ser uma pessoa melhor não só em termos acadêmicos. Obrigada de verdade Ricardo, pelo respeito e empatia com que me acolheu mesmo em meio às nossas diferenças pessoais e pelos momentos de risada que facilitaram tudo.

Agradeço pela leitura dedicada da banca examinadora, de qualificação e defesa, formada por excelentes profissionais e em sua maioria, amigos, Claudette, Cristina Joly, Acácia, Alessandra, Carlos Nunes, Jacob Laros e Robert Evan Verhine.

Ao Inep/Capes/Observatório da Educação, pelo apoio financeiro para a realização desse trabalho.

Aos demais professores da Pós-Graduação da USF, especialmente, Ana Paula, Anna Elisa e Fermino, que ao longo das disciplinas e contatos tanto me ensinaram.

Aos meus grandes amigos ‘adquiridos’ via USF, mas que espero ter por perto ao longo da minha vida, Daniel, Fernanda e Sanyo – vocês tornaram todo o percurso mais sutil com o apoio, amizade, carinho e muitas piadas, é claro. Também de coração gostaria de agradecer pela amizade de pessoas tão especiais, competentes e engraçadas como Lucas, Fabiano, Fernando, Monalisa, Maurício, Maria de Fátima, Andréia, Priscila, Maiana, Samantha, Karine, Adilson, Aleteia, Mauro Salviatti, Maria Eugênia, Ana Vizelli, Claudia Arcanjo, entre tantos outros com quem convivi ao longo desses anos de graduação, mestrado e doutorado na USF. Além desses, dedico carinho especial às funcionárias e atendentes da USF, alunos e responsáveis das instituições fontes de pesquisa, alunos da

graduação da USF e FIEL, colegas do CEFAS, pacientes da clínica, colegas e experiências adquiridas em disciplinas e atendimentos na Unicamp, entre tantas outras.

Bem, são tantas as experiências e pessoas que contribuíram para quem me tornei, que chego praticamente ao final com receio de ser injusta, mas sei que com certeza foram citados todos, direta ou indiretamente...

Resumo

Silva, M. C. R. (2011). *Validade do ENADE para avaliação da qualidade dos cursos: modelo de rasch e multinível*. Tese de Doutorado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba.

A interpretação de resultados de testes educacionais e psicológicos tem sido alvo de investigações nas últimas décadas e juntamente com o estudo das variáveis associadas ao desempenho acadêmico contribui para a compreensão das diferenças em estudantes do ensino superior e instituições de ensino. O presente estudo se propõe a cooperar para essa área trazendo questionamentos de base psicométrica a respeito da metodologia aplicada, a validade das interpretações e os fundamentos psicológicos associados ao ENADE. Foram utilizados dados de 71.838 ingressantes e 58.428 concluintes do curso de Administração do ENADE em 2006, além de informações sobre as instituições de ensino e o perfil socioeconômico dos respondentes. No geral verificou-se a predominância de um fator geral apesar das diretrizes multidimensionais e infere-se que avaliações educacionais como o ENADE tem como base subjacente principalmente o grau de domínio dos estudantes em conteúdos específicos. De acordo com a classificação proposta pelos juízes, mais da metade dos participantes encontraram-se no nível de competência classificado como baixo. Por fim, após o controle de algumas variáveis relacionadas ao perfil anterior ao Ensino Superior como as notas dos ingressantes, renda e formas de sustento e trabalho, constatou-se que somente 4% da variabilidade do desempenho dos concluintes se deve às instituições de ensino, sendo o restante relacionada às diferenças individuais entre os estudantes. Tais resultados demonstram que existe pouca variabilidade para dizer que o exame avalia a qualidade do curso e diferenças em termos de competência. Sugere-se que sejam realizados outros estudos que usem medidas de controle inicial, além de ponderar essas associações e variações nos demais cursos.

Palavras-chave: Avaliação Educacional, Ensino Superior, habilidades

Abstract

Silva, M. C. R. (2011). *ENADE validity for assessing the quality of the courses: rasch model and multilevel*. Doctoral Dissertation, Psychological Graduation Program, São Francisco University, Itatiba.

The interpretation of results of educational and psychological tests have been the subject of investigations in the last decades and along with the study of variables associated with academic performance contributes to understanding of differences among students in faculty and education institutions. This study aims to cooperate in this area bringing questions with psychometric basis about the methodology, the validity of interpretations and the psychological foundations associated with ENADE. The data used referred to 71,838 first-year and 58,428 senior students of several graduation from ENADE's 2006 Administration course, and information about educational institutions and the socioeconomic profile of respondents. Overall there was a predominance of one general factor despite the multidimensional guidelines and it was deduce that educational assessments as ENADE is based mainly in the students degree of dominance in specific content. According to the classification proposed by the judges, over half of the participants met the level of competence classified as low. Finally, after controlling for several variables related to the previous profile to Higher Education as the scores of first-year students, income and livelihood and work, it was found that only 4% of the variability of the performance of graduates is due to educational institutions and the remainder related to individual differences among students. These results demonstrate that there is little variability to say that the test assesses the quality of the course and differences in terms of competence. It is suggested that further studies be conducted that use measures of initial control, and consider variations in these associations and other courses.

Key-words: Educational assessment, Higher Education, abilities

Resumen

Silva, M. C. R. (2011). *Validez del ENADE para evaluar la calidad de los cursos: modelo de Rasch y multinivel*. Tesis de Doctorado, Postgrado en Psicología de la Universidad San Francisco, Itatiba.

La interpretación de los resultados de las pruebas educativas y de investigación psicológica ha sido objetivo de investigaciones en las últimas décadas y, junto con el estudio de variables asociadas con el rendimiento académico contribuye para la comprensión de las diferencias entre los estudiantes de enseñanza superior e instituciones educativas. Este estudio tiene como objetivo cooperar en esta área, levantando preguntas básicas sobre la metodología psicométrica, la validez de las interpretaciones y los fundamentos psicológicos asociados con el ENADE. Se utilizaron datos de 71.838 estudiantes de primer año y 58.428 estudiantes del final de curso de Administración del ENADE en 2006, e informaciones sobre las instituciones educativas y el perfil socioeconómico de los encuestados. En general se observó el predominio de un factor general, a pesar de las directrices multidimensionales, y está claro que las evaluaciones educativas como ENADE se basan principalmente en el grado de dominio de los estudiantes en el contenido específico. De acuerdo con la clasificación propuesta por los jueces, más de la mitad de los participantes se reunieron en el nivel de competencia clasificados como bajo. Finalmente, después de controlar diversas variables relacionadas con el perfil anterior a la Educación Superior como las notas de estudiantes del primer año, sueldo y medios de vida y de trabajo, se encontró que sólo 4% de la variabilidad del rendimiento de los graduados se debe a las instituciones educativas, siendo que el resto están relacionados con las diferencias individuales entre los estudiantes. Estos resultados demuestran que existe muy poca variabilidad para decir que el examen evalúa la calidad del curso y las diferencias en relación a la competencia. Se sugiere que se hagan nuevos estudios con uso de medidas del control inicial, y considerar las variaciones e asociaciones en otros cursos.

Palabras clave: Evaluación educativa, educación superior, habilidades

SUMÁRIO

Lista de Figuras	xii
Lista de Tabelas	xiii
Lista de Anexos	xv
APRESENTAÇÃO.....	1
1. A AVALIAÇÃO EM LARGA ESCALA NO ENSINO SUPERIOR BRASILEIRO	7
2. CONTRIBUIÇÃO DOS CURSOS PARA O DESEMPENHO DOS ESTUDANTES ...	25
3. SUBSÍDIOS PSICOLÓGICOS PARA O ENTENDIMENTO DAS MEDIDAS DE APRENDIZAGEM.....	49
<i>Habilidade e Competência</i>	51
4. ESTUDO PSICOMÉTRICO DO ENADE.....	65
<i>O construto medido pelas provas</i>	65
<i>A métrica das escalas</i>	67
<i>A medida da qualidade dos cursos</i>	74
5. OBJETIVOS.....	78
<i>Objetivo Geral</i>	78
<i>Objetivos Específicos</i>	78
6. ESTUDO 1: Evidências de validade da estrutura interna da prova do ENADE	80
6.1 Método	80
Participantes	80
Materiais	81
Procedimento	82
6.2 Resultados e Discussão.....	83
7. ESTUDO 2: Compreender o sentido dos níveis de proficiência empregando métodos da Teoria de Resposta ao Item	100
7.1 Método	100

Participantes	100
Materiais	100
Procedimento	101
7.2 <i>Resultados e Discussão</i>	101
8. ESTUDO 3: Investigar as evidências de validade por intermédio da verificação das relações com outras variáveis	117
8.1 <i>MÉTODO</i>	117
Participantes	117
Materiais	117
Procedimento	117
8.2 <i>Resultados e Discussão</i>	118
9. CONSIDERAÇÕES FINAIS	133
10. REFERÊNCIAS	137
11. ANEXOS	166

Lista de Figuras

Figura 1. Baseado no modelo conceituado apresentado por Soares (2005)	26
Figura 2. Estrutura da inteligência segundo o modelo de Cattell-Horn-Carroll.....	57
Figura 3 - Mapa dos itens da prova Formação Geral do ENADE 2006 de Administração .	87
Figura 4 - Mapa dos itens do Componente Específico do ENADE 2006 de Administração	90
Figura 5 - Mapa dos itens do Componente Específico do ENADE 2006 de Administração com critérios	103
Figura 6. Relação entre habilidade e probabilidade de resposta por categoria no item CE19	107
Figura 7. Relação entre habilidade e probabilidade de resposta por categoria no item CE12	108
Figura 8. Relação entre habilidade e probabilidade de resposta por categoria no item CE30	109
Figura 9. Relação entre habilidade e probabilidade de resposta por categoria no CE23 ...	110
Figura 10. Relação entre habilidade e probabilidade de resposta por categoria no CE34 .	111
Figura 11. Distribuição de habilidades no componente específico (concluintes)	119
Figura 12. Ranking das instituições no Componente Específico no modelo nulo	125
Figura 13. Ranking das instituições no Componente Específico após controle.....	128

Lista de Tabelas

Tabela 1 - <i>Distribuição de estudantes selecionados no ENADE 2006 de Administração ...</i>	80
Tabela 2 - <i>Distribuição de estudantes no ENADE 2006 de Administração por organização acadêmica.....</i>	81
Tabela 3 - <i>Estatísticas gerais dos estudantes e itens no ENADE 2006 de Administração segundo o modelo de Rasch na parte de Formação Geral</i>	84
Tabela 4 - <i>Estatísticas dos itens em Formação Geral no ENADE 2006 de Administração após calibração pelo modelo de Rasch</i>	85
Tabela 5 - <i>Probabilidade de acerto esperada</i>	86
Tabela 6 - <i>Estatísticas gerais dos estudantes e itens no ENADE 2006 de Administração segundo o modelo de Rasch na parte de Componente Específico</i>	88
Tabela 7 - <i>Estatísticas no Componente Específico no ENADE 2006 de Administração após calibração pelo modelo de Rasch.....</i>	89
Tabela 8 - <i>Distribuição fatorial e de dificuldade dos principais itens nos contrastes</i>	93
Tabela 9- <i>Conteúdos predominantes, habilidades e competências dos itens do contraste positivo (segunda variável latente)</i>	95
Tabela 10- <i>Conteúdos predominantes, habilidades e competências dos itens do contraste negativo (terceira variável latente)</i>	96
Tabela 11- <i>Distribuição dos estudantes segundo critérios e momento do curso</i>	112
Tabela 12 - <i>Distribuição de estudantes no CE por faixa de competência, momento e conceito do curso.....</i>	113
Tabela 13 - <i>Distribuição de estudantes no CE por faixa de competência, momento e categoria administrativa</i>	114

Tabela 14 - <i>Distribuição de estudantes no CE por faixa de competência, momento e organização acadêmica</i>	115
Tabela 15 - <i>Análise de variância fatorial múltipla para testar diferenças de habilidade entre os estudantes segundo o conceito e o momento do curso</i>	116
Tabela 16 - <i>Estatísticas dos estudantes no CE por momento do curso</i>	118
Tabela 17- <i>Escalas das variáveis explicativas do modelo de regressão multinível</i>	122
Tabela 18 - <i>Modelo sem variáveis explicativas e após a inserção das variáveis de controle</i>	129
Tabela 19 - <i>Diferenças de desempenho dos estudantes por conceito antes e após controle de variáveis externas</i>	131
Tabela 20 - <i>Habilidades Aferidas nas questões objetivas da FG do ENADE 2006 de Administração</i>	169
Tabela 21 - <i>Habilidades Aferidas nas questões objetivas do CE do ENADE 2006 de Administração</i>	169
Tabela 22 - <i>Habilidades Aferidas nas questões dissertativas do CE de Administração</i>	174
Tabela 23 - <i>Categorias de resposta nos itens de Componente Específico</i>	178

Lista de Anexos

Anexo 1 – <i>Aprovação do Comitê de Ética</i>	166
Anexo 2 - <i>Objetivos, diretrizes e conteúdos propostos na avaliação do curso de Administração</i>	167
Anexo 3 - <i>Análise dos itens do Componente Específico</i>	178
Anexo 4 - <i>Itens e curvas referentes à habilidade das pessoas e probabilidade de resposta por categoria nos itens de FG e CE</i>	183

APRESENTAÇÃO

A avaliação educacional, especialmente a partir dos anos 90, passou a ser usada, no contexto brasileiro, como tentativa de encontrar um caminho para a solução de alguns problemas educacionais mais emergentes, a fim de que os processos avaliativos influenciassem, entre outros resultados, a elevação dos padrões de desempenho. No entanto, para que isso ocorra é necessária uma reflexão sobre as avaliações operacionalizadas nos vários níveis do sistema educacional, especialmente nas avaliações em larga escala de natureza amostral e supostamente consideradas representativas em termos estatísticos (Vianna, 2003). Devido ao impacto dessas avaliações nos sistemas educativos, bem como nas políticas públicas, é importante verificar quais fatores estão associados ao desempenho acadêmico, bem como a adequação dos sistemas de avaliação e das interpretações das medidas.

No que se refere à avaliação em larga escala no âmbito educacional brasileiro, existem algumas formas nacionais de avaliação das habilidades e competências estudantis, cabendo destaque ao Sistema de Avaliação do Ensino Básico (SAEB), Exame Nacional do Ensino Médio (ENEM) e o Exame Nacional de Desempenho do Estudante (ENADE). Esse último se refere à avaliação do conhecimento adquirido ao longo do curso universitário e é o foco do presente estudo.

No contexto atual de avaliação do ensino superior verifica-se que os estudantes passaram a ser o ponto central da balança das avaliações, visto que suas respostas aos exames nacionais são enfatizadas como medidas de qualidade das instituições e de seus cursos (Leite, 2010). Embora isso ocorra na prática, entende-se que tais apontamentos da autora não são tão explícitos em termos dos discursos em prol das avaliações. Nesse sentido, cabe destaque as colocações dessa mesma autora de que as avaliações

consubstanciadas em exames nacionais estão a contribuir para “formatar” uma subjetivação capitalista do jovem que chega à universidade e, contribuem para o mesmo propósito, os apelos consumistas do marketing das instituições e o estímulo aos procedimentos individuais e produtivistas. Tais questionamentos feitos pela autora enfatizam o caráter competitivo que tem se estabelecido e sido divulgado na mídia, especialmente quando vinculado às instituições privadas, em que o desempenho no ENADE é divulgado como indicativo de qualidade, vindo a repousar em seus resultados o valor de mercado de um curso de graduação. Assim, também se verifica uma tentativa de classificação (*rankings*) das instituições embasada em critérios de desempenho individual ou decorrentes desses resultados, apesar de a avaliação da qualidade englobar idealmente outros indicadores das instituições e o ENADE ocupar somente uma parcela dessa atribuição.

A idéia de ordenar as instituições de ensino por níveis de qualidade, que em inglês se denomina “ranking”, é hoje bastante difundida internacionalmente, e também tem sido adotada no Brasil desde a década de 90, tanto para a educação básica quanto para a educação superior. A distinção, aqui, é se a avaliação inclui um critério sobre o que é considerado satisfatório ou adequado ou se, simplesmente, os resultados são ordenados em relação à distribuição da população. Poucos discordariam hoje que os cursos e instituições devem ser avaliados, e que estas informações devem ser utilizadas tanto pelas próprias instituições, para melhorar continuamente seu trabalho, quanto pelas pessoas e instituições que com elas se relacionam – secretarias de educação, estudantes, empregadores, financiadores públicos e privados. As dúvidas são como avaliar, como difundir estes resultados, e como fazer uso destas avaliações (Schwartzman, 2010). Goldstein (2001) ressalta que muitas das limitações graves associadas ao uso de rankings escolares surgem porque sua publicação oficial incentiva as pessoas a levá-las muito literalmente. Para esse

autor, tal funcionamento, no mínimo, sugere que uma tarefa do governo é garantir a proeminência de advertências necessárias sobre a interpretação desses resultados.

Nesse contexto, um aspecto essencial a ser discutido refere-se à questão da validade da prova (Primi, 2006b) isto é, em que medida o ENADE atingiu os objetivos definidos pelos planos iniciais da prova. Nesse sentido, é preciso então avaliar, dentre outras questões, a validade de tais interpretações ao discutir os procedimentos metodológicos utilizados; o construto psicológico que embasa tais investigações; quais os fatores que influenciam na aprendizagem dos alunos e como realizar interpretações mais qualitativas a partir dos resultados propiciados por avaliações de desempenho, como o ENADE; pontos estes fundamentais no presente trabalho. Especificamente em relação às questões metodológicas, questiona-se aqui essencialmente a validade em utilizar o desempenho dos alunos como representativo da qualidade dos cursos e trabalho das instituições de ensino, visto que dessa maneira observa-se a aprendizagem do estudante (nível 1) e agrega-se essa para avaliar os cursos (nível 2). Dessa maneira, as questões de validade se ampliam e ficam ainda mais complexas; além de outros questionamentos igualmente importantes e interligados a respeito das inferências feitas por medidas não longitudinais, a quantidade de itens X habilidades e competências que se quer avaliar (Primi, 2006b; Primi & cols., 2009a, 2009b, Primi & cols., no prelo; Primi, Hutz & Silva, no prelo), entre outros que serão apontados mais detalhadamente ao longo da discussão teórica.

Em geral, apesar de algumas diferenças entre os pesquisadores a respeito das variáveis extra-acadêmicas mais relevantes, principalmente em termos hierárquicos, há um consenso de que a contribuição da instituição no desempenho acadêmico dos estudantes está condicionada a diversas características extra-acadêmicas, tais como: características sociais, econômicas e culturais da família e do ambiente cotidiano; o conhecimento que os

estudantes já têm quando entram na instituição; além do potencial para aprender e outras características cognitivas individuais associadas (Brandão, 2000; César & Soares, 2001; Ferrão & Fernandes, 2003; Formiga, 2004; Soares, 2004). Em geral, as pesquisas apontam que apesar de haver muitos fatores que possam explicar as diferenças de desempenho dos estudantes nas várias instituições, as variáveis extra-escolares elucidam a maior parte da variação observada em detrimento de fatores estritamente acadêmicos e institucionais.

Como parte da metodologia para diferenciação entre as instituições, essas pesquisas têm apontado para a necessidade de inclusão, como medida de controle do desempenho acadêmico, o perfil do estudante ao entrar na instituição, além de um indicador do nível socioeconômico. Entende-se que esses pontos são complexos pela necessidade de tornar as instituições de ensino estatisticamente iguais de forma a permitir comparações como se essas tivessem tido os mesmos estudantes como ingressantes. Além disso, envolvem temas que vão além da consideração do perfil dos estudantes intra-escola, mas também a longo prazo, a introdução de mudanças no processo de ensino-aprendizagem.

Ainda, na literatura educacional e psicológica, muitos pontos têm sido levantados a respeito das medidas associadas ao desempenho acadêmico. Muitos autores ressaltam a importância da profundidade dos conhecimentos adquiridos no percurso acadêmico (Horn & Cattell 1966; McGrew & Flanagan, 1998). Outros estudos têm enfatizado, além dos processos mentais atrelados, a relevância da mensuração de conhecimentos em domínios especializados e da consideração de experiências adequadas de aprendizagem quando se objetiva a diferenciação entre os estudantes adultos (Ackerman, 1996a, 1996b; Ackerman, 2005; Ackerman, Beier & Bowen, 2002; Castorina & Faigenbaum, 2002; Cronbach, 1957; Hambrick, 2005; Smith & Marsiske, 1994). Portanto, entende-se no presente estudo ser

relevante discorrer a respeito da contribuição da psicologia da inteligência para a interpretação do construto que é medido nesses exames educacionais.

Dada a importância e a implicação das medidas provenientes das avaliações em larga escala nacionais para a educação e sociedade em geral, alguns pontos têm sido discutidos por pesquisadores da área sobre a adequação desses sistemas de avaliação. Tais questões se referem especialmente a abrangência do construto medido pelas provas (Verhine, Dantas & Soares, 2006), o entendimento sobre o potencial da escola em agregar conhecimentos aos estudantes e os fatores associados ao desempenho escolar (Ferrão & Fernandes, 2003), além da interpretação atribuída aos diferentes níveis de desempenho propostos por tais avaliações (Cronbach, 1996).

O presente estudo se propõe a cooperar para essa área trazendo questionamentos de base psicométrica a respeito da metodologia aplicada, a validade das interpretações e os fundamentos psicológicos associados ao ENADE. Em geral, questiona-se qual a validade do ENADE para medir a qualidade dos cursos, e então dizer que o desempenho em um curso implica em competência, sendo todas essas interpretações feitas por meio de um teste e ao nível do aluno.

Desta forma, o primeiro capítulo versará sobre os mecanismos de avaliação do Ensino Superior Brasileiro e a importância de estudos psicométricos que garantam maior legitimidade as medidas decorrentes dos desempenhos dos estudantes. O segundo capítulo tratará sobre questões metodológicas e variáveis envolvidas nas medidas de avaliações educacionais em larga escala, para assim entender qual a contribuição das instituições para a explicação do desempenho dos estudantes avaliados. No terceiro, serão apresentadas as contribuições psicológicas para o entendimento do construto medido pelo ENADE. O quarto capítulo especificará os três aspectos psicométricos a serem investigados neste

6

estudo. A partir desses embasamentos teóricos, segue-se a especificação metodológica dos três estudos, os resultados alcançados em cada um dos mesmos, as considerações finais, referências e anexos.

1. A AVALIAÇÃO EM LARGA ESCALA NO ENSINO SUPERIOR BRASILEIRO

Os sistemas de avaliação em larga escala de instituições educacionais têm como eixo central a mensuração da qualidade das instituições, que envolve indicadores do quanto cada uma contribui para o desenvolvimento de habilidades acadêmicas, competências profissionais e conhecimento de seus estudantes. Tais medidas são importantes, pois têm implicações não só para a gestão dos sistemas de ensino como para o futuro de setores amplos da sociedade que utilizam seus serviços (Primi, Fernandez & Ziviani, 2003; Soares, 2005).

Souza e Oliveira (2003) debatem sobre a centralidade adquirida pela avaliação educacional na configuração dos sistemas de ensino nestas duas últimas décadas, ressaltando que esta tem sido apresentada e justificada como propulsora de sua "qualidade", substituindo a antiga idéia da "igualdade" como principal objetivo das políticas educacionais. Dentro dessa lógica, a avaliação encerra duas potencialidades bastante funcionais. De um lado, torna-se peça central nos mecanismos de controle, que se deslocam dos processos educacionais para os produtos, transferindo-se o mecanismo de controle das estruturas intermediárias para a ponta, via testagens sistêmicas. De outro, a avaliação legitima "valorações" úteis à indução de procedimentos competitivos entre escolas e sistemas para melhorar pontuações nos *rankings*, definidos basicamente pelos desempenhos em instrumentos de avaliação em larga escala. Segundo os autores, tal competição é garantida pela associação entre desempenho e financiamento, podendo redundar em critérios para alocação de recursos, que incidem, até mesmo, em remunerações diferenciadas dentro de sistemas de ensino. Entende-se, no cenário da avaliação superior

brasileira, que as contribuições desses autores estão mais relacionadas à realidade das instituições privadas, em que se verificam conseqüências dos resultados das avaliações na demanda recebida por essas instituições, ou seja, a procura e escolha das instituições a se cursar estando diretamente motivada pelos resultados divulgados.

Castro (2001) destaca que o desenvolvimento de padrões na educação constitui um tema relativamente novo no debate educacional e, por isso mesmo, tem suscitado crescente interesse e gerado intensa polêmica entre os especialistas e os gestores dos sistemas de ensino. O eixo atual deste debate situa-se em dois temas principais, a questão da igualdade e do impacto sobre a aprendizagem. Segundo essa autora, os debates a respeito da equidade têm aumentado desde que começaram a ser realizadas avaliações nacionais, cujos resultados explicitaram os enormes desequilíbrios existentes entre os Estados e os sistemas de ensino. E, portanto, a simples ponderação sobre o estabelecimento de padrões mínimos, que passariam a referenciar o desempenho dos sistemas de ensino, remete inevitavelmente à discussão sobre o problema da igualdade.

No geral, assume-se que as informações fornecidas por avaliações educacionais podem ser úteis para a melhoria dos sistemas, uma vez que permitem uma estimativa dos resultados obtidos com processos interventivos, avaliando se são desejados ou não, além de permitirem um monitoramento constante desses processos, captando seus efeitos positivos e negativos (Sanders & Horn, 1995; Soares, Cesar & Mambrini, 2001; Jesus, 2004). Nesses termos, pode-se entender que a busca primordial deve ser em otimizar os processos escolares, de tal forma que o progresso acadêmico do estudante e seu desempenho escolar sejam sempre positivos.

Segundo Andrade, Tavares e Valle (2000), a avaliação possui um importante papel nas políticas de Educação Superior, uma vez que os custos do ensino superior se tornam

cada vez mais elevados e essas informações são indispensáveis à prestação de contas da qualidade e da amplitude dos serviços que essas instituições prestam à sociedade, em ensino, pesquisa e extensão. Uma das preocupações existentes refere-se à metodologia utilizada para analisar os dados obtidos por estas avaliações e a busca por procedimentos precisos que permitam o acompanhamento do progresso do conhecimento adquirido pelo estudante ao longo do tempo.

Especialmente no que diz respeito ao sistema educacional brasileiro, nos últimos dez anos observa-se que o Ministério da Educação (MEC) tem colocado a avaliação como um dos alvos importantes de suas políticas (MEC, 2005). A partir da década de 90, vários sistemas de avaliação foram propostos, voltados para o ensino fundamental (Sistema de Avaliação do Ensino Básico – SAEB), médio (Exame Nacional do Ensino Médio – ENEM) e superior (Exame Nacional de Cursos – ENC e, recentemente, o ENADE).

O ENC (mais conhecido como Provão) sofreu críticas por membros da comunidade acadêmica e especialistas em avaliação, principalmente em relação à estrutura e formulação das questões (com pouca especificidade, com conteúdo muito generalizado ou mal formulado) quanto à metodologia utilizada para comparação entre cursos (uso de um sistema com referência à norma e da média como representante do desempenho dos cursos), e por buscar caracterizar o desempenho somente de formandos (Castro, 2001; Landeira-Fernandez & Primi, 2002; Primi, Landeira-Fernandez & Ziviani, 2003; Soares, Ribeiro & Castro, 2001; Verhine & Dantas, 2005; Verhine, Dantas & Soares, 2006).

A apreciação de Belloni (1996, p. 11) transcrita a seguir ilustra o conteúdo principal das críticas proferidas:

Trata-se de uma simples medição de resultado final, um procedimento tradicional e insuficiente de avaliação, pois não considera fatores do processo de

aprendizagem e das condições institucionais. Não oferece elementos para a melhoria da instituição e do sistema, pois não identifica as causas das dificuldades; não beneficia o estudante, pois nada poderá ser feito retroativamente. Não corresponde, portanto, ao objetivo fundamental da avaliação qual seja o de melhoria e aperfeiçoamento das instituições ou do sistema. Além disso, pode vir a estimular o aparecimento de um comércio de cursinhos preparatórios, pois o desempenho da prova poderá se transformar em critério de diferenciação entre portadores de um mesmo diploma.

Convém avaliar se tais questionamentos apontados por Belloni (1996) também ainda podem estar atrelados ao ENADE. De forma geral, pode-se considerar que críticas como essas foram propulsoras de mudanças na avaliação superior brasileira. Barreyro (2004) destaca o surgimento do Sistema Nacional de Avaliação do Ensino Superior (SINAES) com objetivo de avaliar as instituições, os cursos de graduação e o desempenho dos estudantes. Segundo esse autor, trata-se do fruto de trabalho da Comissão Especial de Avaliação da Educação Superior (CEA), integrada por acadêmicos e especialistas da temática da avaliação e da Educação Superior, além de membros do ministério da Educação e representantes da União Nacional dos Estudantes. De acordo com Limana e Brito (2005), o SINAES é parte importante de uma política nacional sobre a educação superior e enfatiza os dinamismos dos processos e relações entre as dimensões externas e internas, particular e global, quantitativo e qualitativo e os diversos objetos e objetivos da avaliação.

Criado pela Lei n° 10.861, de 14 de abril de 2004, o SINAES é formado por três componentes principais: a avaliação das instituições, dos cursos e do desempenho dos estudantes. Esse sistema avalia todos os aspectos que giram em torno desses três eixos: o ensino, a pesquisa, a extensão, a responsabilidade social, o desempenho dos estudantes, a

gestão da instituição, o corpo docente, as instalações e vários outros aspectos. Ele possui uma série de instrumentos complementares: auto-avaliação, avaliação externa, ENADE, avaliação dos cursos de graduação e instrumentos de informação (censo e cadastro). Os resultados das avaliações possibilitam traçar um panorama da qualidade dos cursos e instituições de educação superior no País (MEC, 2004).

Os processos avaliativos são coordenados e supervisionados pela Comissão Nacional de Avaliação da Educação Superior (CONAES) e a operacionalização é de responsabilidade do Inep. As informações obtidas com o SINAES são utilizadas pelas IES, para orientação da sua eficácia institucional e efetividade acadêmica e social; pelos órgãos governamentais para orientar políticas públicas e pelos estudantes, pais de estudantes, instituições acadêmicas, e o público em geral, para orientar suas decisões quanto à realidade dos cursos e das instituições.

Diante disso, o ENADE é parte dos instrumentos utilizados pelo SINAES e foi formulado para tentar superar tais limitações do exame anterior. O exame tem como objetivo verificar as competências e habilidades básicas das áreas (como reconhecimento e definição de problemas), os conhecimentos sobre conteúdos básicos e profissionalizantes e verificar o desempenho em questões transdisciplinares, envolvendo o conhecimento mais geral. É examinado o desempenho de estudantes ingressantes e concluintes com a pretensão de analisar as mudanças, os ganhos do estudante ao longo de sua trajetória na Instituição (MEC, 2004a).

A prova do ENADE é composta de quarenta questões, das quais dez se propõem a avaliar a formação geral e trinta questões são específicas de cada área, ambos os formatos contendo questões discursivas e objetivas de múltipla escolha. Nas questões da prova busca-se também obter indícios relativos à capacidade do estudante para analisar, sintetizar,

criticar, deduzir, construir hipóteses, estabelecer relações, fazer comparações, detectar contradições, decidir e organizar as idéias. No componente de Formação Geral (FG), dentro dos limites possíveis, é investigada a formação de um profissional ético, competente e comprometido com a sociedade em que vive, e são utilizadas situações-problema, estudos de caso, simulações e interpretação de textos e imagens. As questões discursivas buscam investigar, além do conteúdo específico, aspectos como a clareza, a coerência, a coesão, as estratégias argumentativas, a utilização de vocabulário adequado e a correção gramatical do texto (MEC, 2005).

O Componente Específico (CE) de cada carreira é elaborado com base nas Diretrizes Curriculares, aprovadas pelo Conselho Nacional de Educação e no perfil profissional de cada carreira, contemplando os saberes fundamentais exigidos em cada área profissional. Assim, o componente específico busca contemplar a especificidade de cada área, tanto no domínio dos conhecimentos quanto nas habilidades esperadas para o perfil profissional (MEC, 2005).

No Anexo 2 são apresentados os objetivos, habilidades e competências, assim como os conteúdos a serem avaliados no curso selecionado para esse estudo. Cabe ressaltar que, além das especificações das habilidades e competências que basearam a formulação de cada uma das provas dos cursos, no Relatório Técnico das áreas avaliadas também podem ser analisadas as diretrizes para cada uma das questões. Essas informações são muito relevantes, visto que permitem, a partir da matriz de avaliação proposta, realizar uma análise de conteúdo com o auxílio de especialistas da área a fim de verificar a abrangência dessas no exame realizado. Tal proposta de avaliação para o presente estudo será mais bem detalhada em outro momento mais apropriado.

A partir das Diretrizes Curriculares Nacionais e pelas definições da Comissão Assessora de Avaliação nas diferentes áreas, entende-se que o ensino possa promover ao profissional um domínio básico de conhecimentos e a capacidade de utilizá-los em diferentes contextos que demandam a investigação, análise, avaliação, prevenção e atuação em situações definidas (MEC, 2005). O que se torna importante saber nessa perspectiva é, em que medida é possível diferenciar um estudante mais competente de outro menos, e quais capacidades intelectuais estão envolvidas neste processo.

Em geral, entende-se que o ENADE busca aferir o desempenho do estudante e, seus resultados, podem mostrar se o estudante teve ou não desenvolvimento nas suas competências, se é capaz de trabalhar os conteúdos aprendidos e sua capacidade para usar esse conhecimento no seu desenvolvimento (MEC, 2004b). Segundo Limana e Brito (2005) o ENADE permite um acompanhamento dos estudantes de forma a entender melhor o processo, e envolve a estimativa do potencial do ingressante e a aquisição de competências pelo concluinte. Segundo esse conceito, a avaliação é dinâmica e permite reunir informações sobre as mudanças de desempenho dos estudantes durante a trajetória na universidade. O potencial de aprendizagem é entendido como o conjunto de conceitos e procedimentos considerados essenciais para o domínio da área. Diante das considerações desses autores, cabe avaliar mais criticamente se na prática o ENADE conseguiu avançar em todos esses pontos considerados frágeis no ENC, questionamentos que se espera poder contribuir com o presente trabalho.

Brito (2006) discorre que a passagem do estudante pela instituição produz mudança e esta não é apenas na quantidade de conhecimento, mas também no domínio das tarefas específicas da área, na maneira como ele usa as diferentes competências que desenvolve e a

capacidade para estender esta competência a novos domínios; sendo capaz de transferir significativamente os conteúdos em diferentes contextos, principalmente profissionais.

Como o ENADE também se propõe a avaliar as competências profissionais dos estudantes, cabe destacar algumas das considerações realizadas por Gondim (2002), no que se refere à abrangência da formação como promotora de competências básicas para atuação de estudantes de diferentes áreas. Essa autora conduziu uma pesquisa qualitativa com o objetivo de investigar as expectativas de inserção futura no mercado de trabalho e da percepção das competências profissionais necessárias para atuação em estudantes universitários em fase de conclusão de curso. A autora concluiu que o estudante consegue identificar um perfil geral que é transmitido a respeito das exigências do mercado de trabalho, mas não consegue identificar de que modo este perfil esteja ajustado às habilidades e competências próprias de sua área de atuação, e especificamente vinculados em sua formação.

Outra problemática referente às implicações das avaliações educacionais em larga escala, como a proposta do ENADE, foi debatida por Souza e Oliveira (2003). De acordo com os autores, tais propostas têm uma intencionalidade de proceder a uma avaliação mais abrangente das instituições, porém ficam indagações quanto ao uso de seus resultados, especialmente como referência de qualidade de um dado curso e, por decorrência, da instituição. Afora o incentivo à competitividade, provocado pelo estabelecimento de *ranking* de instituições, cabe indagações sobre possíveis usos dos resultados, seja no estabelecimento de critérios para alocação de recursos financeiros entre as instituições, tendo em conta seu bom ou mau desempenho, seja na política de financiamento do ensino superior.

Além dessas questões referentes à abrangência e promoção da formação, Verhine,

Dantas e Soares (2006) fizeram uma análise comparativa entre as propostas avaliativas do Provão e do ENADE. Estes autores concluíram que existe uma similaridade dos processos técnicos das duas abordagens, enquanto que as diferenças se apresentam mais no plano conceitual, uma vez que o ENADE se propõe a apresentar um indicador de diferença de desempenho entre ingressantes e concluintes. No entanto, os autores ressaltam que o novo exame não conseguiu, até o momento, resolver vários dos problemas identificados na abordagem anterior, mas que algumas das propostas, se implementadas, constituirão um avanço real para a utilização dos seus resultados. Uma das críticas centrais deste exame se refere à profundidade em que as habilidades e competências podem ser avaliadas, comparativamente ao número reduzido de itens.

Nesse sentido, embora o ENADE tenha avançado bastante em relação ao Provão, ele também está sujeito a muitos questionamentos do ponto de vista metodológico como já apontado por Verhine, Dantas e Soares (2006) e também pelos autores Primi e cols. (2009a, 2009b). Primi e cols. (2009b) centrou-se em discutir questões metodológicas relacionadas com a limitação do desenho transversal assumido no ENADE e as suposições em que está fundamentado que desta forma, colocam em dúvida as interpretações que são feitas. Os autores apontam que demais questionamentos são mais amplos, não só circunscritos ao ENADE, mas aos estudos de valor agregado, especialmente sobre a dificuldade em isolar os efeitos das instituições na aprendizagem do aluno. Primi e cols. (2009a) discutiram as vantagens da aplicação da Teoria de Resposta ao Item na avaliação criterial e substancial dos itens envolvidos no ENADE 2006 de Psicologia. Nesta pesquisa, utilizou-se a TRI para a determinação de pontos de corte, formando grupos de competências requeridas para a resolução de itens. A escala de habilidades dos estudantes foi dividida em quatro níveis: insuficiente, fraca, competência mínima e competência

superiora. Os resultados demonstraram que, de maneira geral, os estudantes concluintes concentram-se na competência mínima, enquanto que os ingressantes distribuem-se, em sua maioria, em fraca e competência mínima. Também foram analisadas as distribuições de acordo com o conceito da instituição. Considera-se que esse sistema de interpretação criterial do ENADE auxilia mais amplamente a análise da prova e dos estudantes.

De forma complementar, Primi, Hutz e Silva (no prelo) estudaram também a prova do ENADE 2006 de Psicologia e destacam que outro aspecto questionável do ENADE tem relação com o conteúdo das provas e a sua validade em avaliar as competências profissionais e habilidades acadêmicas conforme planejado. Mais especificamente, no que se refere aos parâmetros referenciais para sua interpretação, o formato e a quantidade de itens da prova.

Também a partir dos dados do ENADE 2006 desse mesmo curso, Primi, Carvalho, Miguel e Silva (no prelo) tentaram detectar Itens com Funcionamento Diferencial (DIF), isto é, itens com problema de equivalência ao medir ingressantes e concluintes e estudantes de instituições públicas e privadas. Empregou-se a análise de Rasch (TRI) e regressão logística para se detectar o DIF e encontrou-se que 11 dos 30 itens que compunham a prova apresentaram DIF. Os autores apontaram que a presença de DIF significa que, mesmo considerando pessoas com o mesmo nível de habilidade a partir do escore global na prova, estudantes concluintes e os de universidade pública tendem a ter vantagem nos itens com DIF. Também discutem alternativas metodológicas para se tentar lidar com esse problema, como a resolução do DIF, e que o julgamento do viés precisa ser feito com base em uma discussão mais ampla do conteúdo e da análise da influência que esses itens têm na medida global obtida a partir das pontuações dos estudantes ao responderem a prova.

Alguns outros estudos têm sido realizados a partir dos dados do ENADE disponíveis pelo INEP. Como exemplo disso, pode ser citado o trabalho de Silva, Bartholomeu, Pires e Primi (2010) que fizeram um estudo a partir de uma prova equalizada com o ENADE 2006 de Administração por meio da TRI e os resultados possibilitaram a identificação do nível de habilidades e competências dos alunos avaliados por essa prova equalizada em comparação ao que é proposto pelo ENADE. Tal metodologia permite avaliar um grupo específico de estudantes em instituições de ensino superior que ainda não puderam passar pelo ENADE e identificar déficits no aprendizado dos alunos, o que pode conduzir as instituições ao planejamento de programas de ensino que procurem suprir tais necessidades dos alunos, melhorando o sistema de ensino como um todo. Os autores apontam que, em geral, a vantagem desse estudo consiste em possibilitar uma medida, na mesma métrica do ENADE, e que permite a comparação das instituições avaliadas em relação às demais no Brasil num curso específico.

Além de possibilitar a análise comparativa com os resultados do ENADE, Bartholomeu, Silva, Pires e Primi (2010) analisaram outra prova equalizada com o ENADE 2006, do curso de Direito, a fim de realizar uma análise qualitativa dos itens e, assim, avaliar se as categorias explicativas poderiam indicar os fatores associados à dificuldade dos itens (análise dos itens em termos das exigências de conteúdos/habilidades/competências). A análise foi realizada por um grupo de juízes considerando a presença das habilidades e conteúdos nos itens dessa prova, a partir do que foi proposto inicialmente na matriz do ENADE. Os juízes realizaram o estabelecimento de pontos de corte a partir da hierarquia de dificuldade dos itens dada pelo modelo de Rasch (TRI). Foram apontadas dificuldades nesse processo visto que na prova é proposta a aferição de uma quantidade muito grande de conteúdos e habilidades em um único item. Em geral, os

autores apontaram que essa análise mais criterial dos resultados e conteúdo dos itens em cada um dos níveis de corte possibilitou uma comparação mais adequada dos estudantes em termos de seus níveis de proficiência além de desvendar algumas características psicológicas e institucionais associados aos mesmos.

Entende-se que apesar do ENADE se constituir um avanço em relação ao ENC, em geral existem ainda muitas críticas a esse exame e as interpretações dele decorrentes. Os pontos centrais dessas, apontadas ao longo do texto, estão relacionados ao uso de uma metodologia transversal; a quantidade e formato dos itens diante da especificação de conteúdos, habilidades e competências proposta nas diretrizes do exame; a mensuração de ingressantes e concluintes a partir da mesma prova; a falta de uma medida de controle inicial do conhecimento do estudante antes da entrada na instituição; a inexistência de estudos consistentes de valor agregado. Tais questões são essenciais quando se objetiva entender o processo, o potencial do ingressante e a aquisição de competências pelo concluinte, além das mudanças de desempenho dos estudantes durante a trajetória, como apontado anteriormente por Limana e Brito (2005).

Portanto, tais críticas confirmam a necessidade de uma avaliação mais criteriosa a respeito dos procedimentos adotados para criação de instrumentos de avaliação educacional e, mais precisamente, a respeito das implicações das interpretações realizadas a partir de tais medidas, temáticas que serão mais pormenorizadas nos capítulos seguintes. Especificamente quanto ao ENADE, entende-se ser essencial uma análise criteriosa feita por grupos de especialistas sobre a pertinência e abrangência das questões que compõe a mesma, garantindo a representatividade do domínio. Mas somente esse procedimento não garante a legitimidade das muitas interpretações pretendidas, pois se entende que os instrumentos utilizados nesses processos de avaliação do estudante devem possuir certas

características que confirmam validade e precisão aos resultados. Esse tipo de estudo oferece uma base científica às provas e aos resultados conseguidos, pondo em evidência, possíveis vantagens e falhas metodológicas na construção desses instrumentos e garantindo instrumentos com maior credibilidade para serem utilizados num processo de avaliação do ensino.

Deste modo, para que a avaliação seja confiável e permita generalizações de resultados amostrais é fundamental que sejam utilizados instrumentos válidos e fidedignos. Embora as provas ou testes educacionais não tenham o mesmo significado que testes psicológicos, com o progresso alcançado na psicometria, diversas técnicas de construção de testes são utilizadas na construção de exames escolares.

Do ponto de vista psicométrico, para que uma avaliação se efetive com êxito, além da elaboração adequada das provas utilizadas, é importante dar atenção à qualidade científica dessas provas. É primordial que se investigue a validade dos instrumentos utilizados na avaliação (Urbina, 2007). Esta é estabelecida usando como referência o uso específico para a qual a prova se propõe, ou seja, relaciona-se com aquilo que pretende medir. Os métodos de validação são vários, sendo denominados de acordo com a ênfase que se pretende avaliar (American Educational Research Association - AERA, American Psychological Association - APA, National Council on Measurement in Education - NCME, 1999; Urbina, 2004).

De acordo com essa visão, a validade é concebida como um conceito único e passa-se a procurar nos testes por evidências de validade e não mais por tipo de validade. Assim, o teste deixa de ser válido e sim, passa a possuir evidências de validade para interpretação de um determinado tipo de resultado e, a proposta de interpretação passa a referir-se ao construto psicológico ou aos conceitos que o teste pretende avaliar. Assim, pode-se dizer

que um teste pode ser válido em determinadas circunstâncias que descrevem um tipo específico de uso, para as quais os dados teóricos ou empíricos encontraram sustentação para suas interpretações (AERA, APA, NCME, 1999).

A AERA, APA e NCME (1999) categorizaram os procedimentos utilizados para a investigação da validade dos testes, sejam eles: *evidência baseada no conteúdo*, na qual se investiga se os itens do teste se organizam em amostras abrangentes do domínio que se pretende avaliar; *evidência baseada nas relações com outras variáveis*, na qual são buscados padrões de correlação entre escores do teste e outras variáveis que meçam o mesmo construto ou construtos relacionados e com variáveis que meçam construtos diferentes; *evidência baseada na estrutura interna*, na qual busca-se dados sobre a estrutura das correlações entre itens avaliando o mesmo construto e ainda sobre as correlações entre subtestes avaliando construtos similares; *evidência baseada no processo de resposta*, na qual busca-se dados sobre os processos mentais envolvidos na realização das tarefas propostas pelo teste; e *evidência baseada nas conseqüências da testagem*, na qual se examina conseqüências sociais intencionais e não intencionais do uso do teste para verificar se sua utilização está tendo os efeitos desejados de acordo com os objetivos para o qual foi criado. A investigação dessas evidências de validade é essencial na testagem psicológica e educacional, como também apontado pelos autores Nunes e Primi (2010), Primi, Muniz e Nunes (2009) e Primi e Nunes (2010).

Outra qualidade psicométrica considerada como importante nos instrumentos é a fidedignidade ou precisão, que avalia a consistência de uma medida quando aplicada em um determinado grupo (Urbina, 2007). Além disso, segundo a AERA, APA e NCME (1999) é comum encontrar nas medidas o pressuposto de que há alguma estabilidade nos comportamentos de indivíduos e/ou grupos. De forma geral, podem ser encontradas três

grandes categorias de mensuração de fidedignidade, sendo a primeira derivada da administração de formas paralelas em sessões de teste independentes (forma alternada); a segunda, a administração do mesmo instrumento em sessões separadas (teste-reteste) e, por último, por intermédio da relação de escores derivados dos itens individuais ou dos itens com o teste em uma única aplicação (consistência interna).

Um dos passos utilizados na construção e avaliação psicométrica de instrumentos psicológicos e educacionais envolve a análise de itens que compõem o instrumento, realizada pela Teoria Clássica dos Testes (TCT) ou pela Teoria de Resposta ao Item (TRI). A TRI foi desenvolvida a partir dos estudos iniciais de Lord (1952/1953) nos Estados Unidos e de Rasch (1960), na Holanda. A TRI é um conjunto de modelos onde a probabilidade de resposta a um item é modelada como função da proficiência do aluno (variável não observável) e de parâmetros (que expressam certas propriedades) do item. Quanto maior a proficiência, maior a probabilidade de o aluno acertar o item.

Essa teoria não entra em contradição com os princípios da psicometria clássica e traz uma nova proposta estatística de análise centrada nos itens que supera as limitações da Teoria Clássica (como a não dependência do conjunto de itens que compõem o teste), além de apresentar novos recursos tecnológicos para a avaliação psicológica e educacional (Andrade, Tavares & Valle, 2000; Anastasi & Urbina, 2000; Primi, 1998; Urbina, 2007).

Geralmente, os testes são compostos de itens que operacionalizam medidas em diferentes níveis de complexidade de uma dimensão latente e quando agregados permitem estimar a habilidade de uma pessoa nessa dimensão. A análise psicométrica dos itens verifica a sua eficácia em medir a suposta dimensão latente. Sob a ótica da TRI, os itens são caracterizados por parâmetros e índices de ajuste os quais permitem verificar sua qualidade. Os principais parâmetros são a dificuldade que se refere ao nível de complexidade do

construto que o item representa e a discriminação, que está associada à correlação do escore no item com o construto medido pelo teste (Primi, Carvalho, Miguel & Silva, no prelo).

Em geral, entende-se que a TRI é um procedimento utilizado quando se deseja estabelecer uma escala de mensuração que independa da amostra, aplicável a indivíduos de níveis variados do construto, ou de conteúdo com níveis muito variados de dificuldade. Uma das vantagens da utilização da TRI é que esta permite análises qualitativas a partir dos resultados brutos do teste, fornecendo assim interpretações mais criteriosas do desempenho dos respondentes submetidos a uma prova e da qualidade dos itens utilizados.

Alguns pontos têm sido levantados na literatura sobre a adequação desta teoria na área de Avaliação Psicológica e Educacional, especialmente no que se refere à dimensionalidade dos traços latentes envolvidos na avaliação e a equalização de diferentes avaliações (Andrade, 2001). Alguns estudos tiveram como enfoque a avaliação das qualidades psicométricas do ENADE com o uso da TRI, como se segue.

Primi (2006b) enviou um projeto a Capes/Observatório da Educação a fim de avaliar a Validade do ENADE para Avaliação da Qualidade dos Cursos de Instituições de Ensino Superior. Após a aprovação do mesmo, destacam-se vários estudos do grupo de professores e estudantes de pós-graduação do Laboratório de Avaliação Psicológica e Educacional (LabAPE) ligado ao Programa de Mestrado e Doutorado em Avaliação Psicológica da Universidade São Francisco, que tem desenvolvido desde 2007 trabalhos a fim de analisar as provas do ENADE por meio da TRI. Dentre os vários estudos desenvolvidos ao longo desses anos, cabe destacar alguns. Oliveira (2006) analisou as propriedades psicométricas do ENADE 2004 do curso de Medicina, sua dimensionalidade e adequação aos modelos da TRI. Primi e Miguel (2006) avaliaram também a prova de Medicina do ENADE 2004 a fim de estudar uma medida de critério para interpretação dos

resultados. Nogueira (2008) aplicou a TRI para avaliar as questões de formação geral do ENADE 2004 e 2005, em especial aquelas que envolvem conceitos estatísticos, visando estimar a proficiência dos estudantes nos conteúdos avaliados e o ajuste dos itens ao modelo de Rasch. Vendramini (2007) analisou as propriedades psicométricas das provas do ENADE 2004 e 2005. Vendramini (2009) realizou a análise dos componentes principais dos resíduos do componente de Formação Geral do ENADE desses mesmos anos pela TRI. Primi e cols., (2009a) e Primi, Hutz e Silva (no prelo), além de analisarem as propriedades psicométricas, utilizaram a TRI para a determinação de pontos de corte da prova de Psicologia do ENADE 2006, formando grupos de competências requeridas para a resolução de itens. Primi, Carvalho, Miguel e Silva (no prelo) avaliaram o Funcionamento Diferencial dos Itens (DIF) a fim de verificar diferenças de desempenho em alguns itens ao medir ingressantes e concluintes, assim como estudantes de instituições públicas e privadas.

Além disso, um dos objetivos do grupo foi a criação e análise de uma prova equalizada com o ENADE, do Programa de Avaliação Continuada (PAC), a fim de analisar, acompanhar e verificar o desempenho dos estudantes da Universidade São Francisco e compará-lo com os demais universitários brasileiros avaliados pelo ENADE (Bartholomeu, Silva, Pires & Primi, 2010; Carvalho & cols., 2007; Dias & cols., 2007; Primi & cols., 2007a, 2007b; Silva & cols., 2007; Silva, Bartholomeu, Pires & Primi, 2010; Vendramini & cols., 2007a, 2007b; Primi & cols., 2008a, 2008b; 2008c; entre outros).

Além das questões específicas centradas na composição da prova propriamente dita, é importante considerar que o desempenho em cursos distintos associa-se ao domínio de conteúdos também distintos, e, assim, pode-se supor que a importância relativa do raciocínio e conhecimento pode mudar se o desempenho em diferentes cursos for analisado separadamente (Primi, Santos & Vendramini, 2002). Essa interpretação psicológica do

exame é essencial neste trabalho e será mais detalhada posteriormente, já que o ENADE se propõe a avaliar os diversos cursos das áreas de Humanas, Exatas e de Saúde, por meio de provas específicas que pretendem analisar as habilidades e competências exclusivas de cada um destes campos conhecimento e por meio de uma prova única para ingressantes e concluintes.

Entende-se que a contribuição psicológica do presente estudo centre-se em discutir, por meio das interpretações propostas pelo ENADE, qual o construto psicológico medido pelas provas e assim entendido como elemento diferenciador entre os estudantes, tendo como base as teorias de inteligência e, por conseqüência, qual o papel dos cursos e das instituições de ensino nesse processo. Além disso, esclarecer, acerca de quais variáveis podem interferir nos variados níveis de desempenho dos estudantes, a partir do ponto de vista qualitativo e da aplicação de métodos avançados de análise de dados. Por fim, dar subsídios para uma interpretação criterial dos resultados obtidos por meio do exame.

Desse modo, o capítulo seguinte tem como proposta central discorrer sobre os posicionamentos teóricos e metodológicos que propõem explicar as contribuições e diferenças de qualidade entre os cursos das diferentes instituições de ensino.

2. CONTRIBUIÇÃO DOS CURSOS PARA O DESEMPENHO DOS ESTUDANTES

A fim de contextualizar a temática da avaliação do sistema escolar Brandão (2000) aponta que essa foi feita historicamente e quase que exclusivamente, com base na comparação dos estudantes das diferentes regiões do país freqüentando os vários níveis e tipos de ensino (regular, técnico, entre outros) nas diversas redes de ensino (federal, estadual, municipal e privado), assim como nas zonas urbanas ou rurais. Esse tipo de avaliação fundamentava-se em uma *abordagem externa*, pois a interpretação do funcionamento do sistema escolar era desenvolvida, indiretamente, pelos efeitos da escola sobre a trajetória dos usuários representados nas faixas de idade e fluxos escolares. Gradativamente, procurou-se aprofundar a interpretação desses fluxos analisando a progressão diferencial nas redes de ensino, segundo as características demográficas e sociais das populações. Essa autora destaca que uma nova direção da pesquisa sobre a seletividade nas e das escolas contribuiu, de forma decisiva, para a compreensão da trama complexa de interações entre os fatores intra e extra-escolares que tecem os processos de escolarização. Mais especificamente, os estudos desses processos permitiram, por exemplo, o aprofundamento do conhecimento do impacto do capital cultural das famílias sobre o desempenho escolar. Ou seja, tal perspectiva da autora tem como foco a verificação das causas e fatores diretamente associados ao desempenho, sejam eles externos ou internos.

Como a avaliação do sistema escolar mostra-se diretamente relacionada ao desempenho e sua avaliação, convém levantar os fatores que determinam o ensino e a aprendizagem. Goldstein e Woodhouse (2008) criticam o empenho de muitos pesquisadores em produzir listas de fatores importantes de forma a simplificar as

complexas relações de causa e efeito associadas ao ensino. Também criticam a tendência predominante entre os pesquisadores de verem a escola como entidade independente, sem interferências externas, quando se sabe que a mesma opera como fator integrante de um sistema social e político, no qual cada escola não só influencia seu meio, mas também é influenciada por ele.

Nesse contexto, a Figura 1 apresenta o modelo elaborado por Soares (2005), fruto da consolidação de vários outros modelos existentes na literatura.

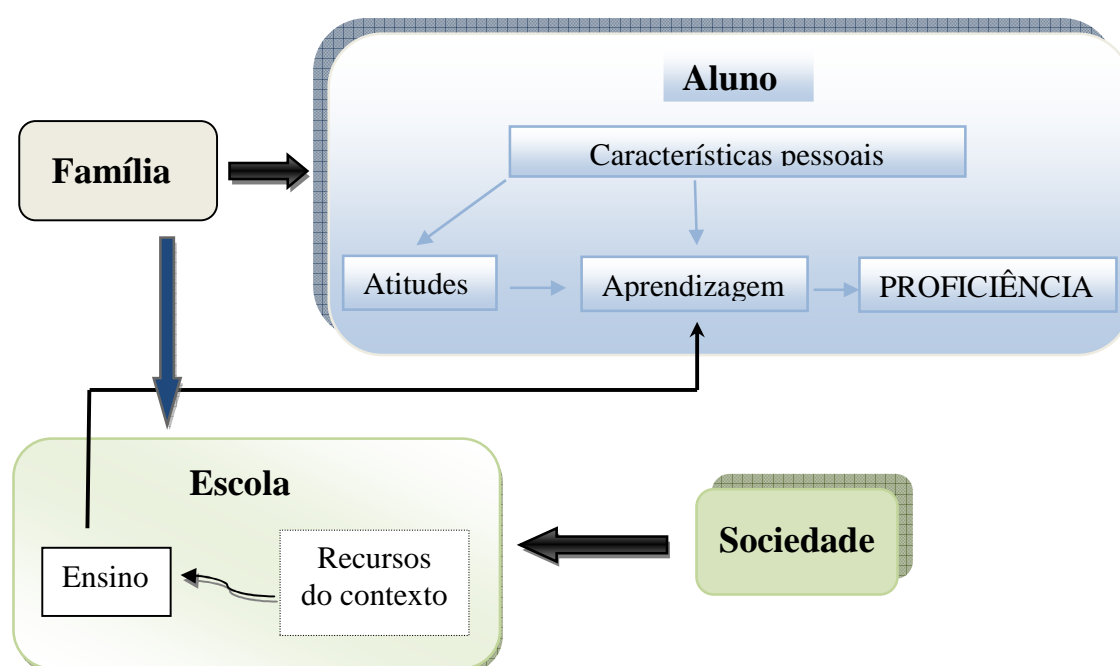


Figura 1. Baseado no modelo conceituado apresentado por Soares (2005)

De acordo com esse autor, os fatores que determinam o desempenho pertencem a três grandes categorias: os associados à estrutura escolar, os associados à família e àqueles relacionados ao próprio estudante (ponto de vista cognitivo). Soares (2004) destaca que pesquisas empíricas conduzidas nos anos de 1950 e 1960 nos Estados Unidos, Inglaterra e França mostraram que fatores extra-escolares explicam mais as desigualdades observadas no desempenho dos estudantes do que fatores intra-escolares. De forma particular,

mostraram que tanto o acesso à educação como os resultados escolares estão associados de maneira forte e direta às características socioeconômicas e culturais dos estudantes.

Ao se analisar o modelo apresentado na Figura 1, constata-se a complexidade do fenômeno desempenho escolar, de forma que os fatores apontados como mais próximos do desempenho do aluno são suas características inatas ou já determinadas por sua história de vida. Segundo esse autor as características pessoais como raça, sexo, saúde, trajetória escolar e talentos inatos explicariam as atitudes em relação à escola, bem como a aprendizagem e proficiências. Além dessas, a escola (que envolve o professor, modo de gestão, outros alunos, recursos físicos, humanos e pedagógicos, entre outros), a família (com seus recursos econômicos, culturais, envolvimento dos pais e modo de estruturação) e a sociedade disputam para melhores ou piores desempenhos de alunos. Assim, para o entendimento completo do desempenho do aluno é necessária uma abordagem multidisciplinar que agregue conhecimentos de várias ciências, dentre elas a psicologia, em muitos momentos subsidiados pela coleta e análise de dados através de técnicas estatísticas apropriadas (Soares, 2007).

O acompanhamento dos progressos escolares remete à questão da eficácia escolar, foco da maior parte dos trabalhos que tratam das avaliações educacionais na atualidade. Madaus, Airasian e Kellagan (2008) apontam para algumas diferenças na utilização da terminologia eficiência e eficácia, e assim, na adequação dessa última nas pesquisas sobre as diferenças de atribuição da escola no desempenho dos estudantes. Para esses autores, a eficiência de um sistema é avaliada pelos seus custos, de forma a estabelecer se os produtos correspondem aos investimentos realizados, enquanto que eficácia está diretamente associada com a qualidade das instituições e, então, na capacidade dessas de produzirem resultados efetivos esperados do ponto de vista social. Nesse sentido, uma escola é tida

como eficaz quando apresenta características que assegurem a efetividade e eficácia de seu ensino produzindo bons resultados escolares em seus estudantes (Soares, Sátyro & Mambrini, 2000). É nesse sentido que as instituições responsáveis pela gestão de políticas públicas têm particular interesse nos fatores que diferenciam as escolas eficazes e não eficazes.

Nesse contexto, cabe destacar diferenças nas denominações que caracterizam as pesquisas em eficácia escolar, em que há pesquisadores que preferem nomear como o estudo do efeito-escola e outros, a identificação de escolas eficazes, como destacam Brooke e Soares (2008):

Por efeito-escola entende-se o quanto um dado estabelecimento escolar, pelas suas políticas e práticas internas, acrescenta ao aprendizado do aluno. Essa definição enfatiza a idéia de que cada escola deve ser analisada a partir dos resultados de seu processo de ensino-aprendizagem e que os fatores associados com os melhores resultados devem ser identificados. Já o termo escola eficaz, principalmente nos primeiros trabalhos, sugere apenas que existem escolas melhores do que outras. Hoje, no entanto, ambos os termos são usados (p. 10).

Na literatura científica, a medida da qualidade das instituições ou da eficácia escolar é baseada em medidas do valor agregado, definido como “uma medida do progresso médio do estudante, durante o período de tempo em que ele é exposto a determinado ambiente educativo, comparativamente com o que obteria em outras escolas da amostra” (Ferrão, 2003, p. 16). Ou seja, o valor agregado corresponderia à diferença entre conhecimento e habilidades do estudante no momento de sua entrada comparativamente ao momento de sua saída (Soares, Ribeiro & Castro, 2001; Jesus & Laros, 2004; Raudenbush, 2004b; Rubin, Stuart, & Zanutto, 2004). Nesse estudo, entende-se como essencial a verificação dessas

medidas por meio de estudos delineados para testar se as interpretações que são feitas a partir dos dados dos instrumentos de medida são plausíveis, assim investigando tanto os instrumentos quanto os delineamentos de coleta de dados.

Assim, os estudos de eficácia escolar ou efeito-escola partem do conceito de valor agregado, termo que teve sua origem nas ciências econômicas e foi amplamente utilizado na política nos anos 90. Esse termo foi criado para expressar a mensuração dos resultados escolares de estudantes e, conseqüentemente, da qualidade da educação oferecida. Goldstein (1997) sugere que a quantificação do valor agregado na educação é baseada na busca de comparações válidas entre as escolas quanto aos resultados obtidos por seus estudantes.

Gray (2008) destaca que a assimilação do conceito de valor agregado se deve à dificuldade de usar dados transversais para descrever a contribuição específica da escola no desempenho dos alunos e a falta de uma medida precisa de controle da influência da aprendizagem anterior do aluno no cálculo do efeito escola. Segundo esse autor, para calcular o valor agregado deve-se tomar por base o progresso de todos os alunos por meio de uma medida de seus respectivos desempenhos na ocasião do ingresso escolar. E, por tomar como ponto de partida o desempenho inicial do mesmo aluno, a medida de valor agregado elimina a necessidade de controle pelo aprendizado anterior à entrada da escola e permite definir a eficácia da escola de acordo com o acréscimo na aprendizagem dos alunos além do que se esperaria ao considerar suas características de origem.

Com base nessa distinção, Raudenbush e Willms (1995) propuseram dois tipos de efeito-escola, a saber, A e B. O primeiro é alcançado estabelecendo a diferença entre o desempenho real do aluno e o que ele teria se frequentasse uma escola típica, sendo necessário o ajuste de um modelo estatístico que controle a caracterização sócio-

demográfica e cultural dos estudantes. Um efeito do tipo A inclui os efeitos das políticas e práticas escolares, da composição de estudantes da escola, fatores sociais e econômicos exógenos, além de quaisquer outros efeitos não medidos associados à determinada escola. Logo, os efeitos tipo A poderiam ser denominados como diferenças ajustadas da escola, porque incluem os fatores que escapam ao controle dos professores e administradores (Willms, 2008).

A sua vez, o efeito Tipo B considera mais a influência das características administrativas, pedagógicas e curriculares da escola. Nesse sentido, é a diferença entre o desempenho real do aluno e seu desempenho esperado caso ele estivesse em uma escola com contexto idêntico, sendo necessário controlar características do aluno e do contexto em que a escola está inserida. Efeitos do tipo B são úteis para responder a questão: “qual o desempenho da escola X em relação a outras escolas com conjunto de alunos similar e contextos social e econômicos também análogos?” Assim sendo, as estimativas do tipo B são obtidas através de modelos estatísticos que incluem variáveis que controlam as influências decorrentes da composição social dos alunos da escola e os efeitos decorrentes de fatores sociais e econômicos que escapam ao controle do sistema escolar (Willms, 2008). Pode-se então dizer que os estudos que envolvem o efeito A calculam o valor agregado controlando essencialmente as diferenças prévias do estudante e os que envolvem o estudo do efeito B levam em conta essa diferença, mas controlam especificamente as variáveis específicas do contexto escolar.

De forma geral, as pesquisas desenvolvidas nessa área comumente empregam o termo efeito-escola (seja A ou B) para designar a explicação do desempenho escolar dos estudantes após o controle de algumas variáveis, denotando assim a eficácia escolar. Muñoz-Repiso e Murillo (2003) indicaram que as linhas de investigação nessa área têm se

baseado essencialmente em estimar a magnitude dos efeitos escolares e análise de suas propriedades científicas (confiabilidade, consistência, dentre outros); e no estudo de fatores escolares da aula e de contexto, com vistas a caracterizar uma escola eficaz.

Assim, a interpretação dos valores agregados pelas escolas depende essencialmente das variáveis selecionadas, que devem, a sua vez, partir de hipóteses aventadas pelo ponto de vista teórico-empírico. Ainda, a confiabilidade das variáveis é raramente indicada nessas pesquisas (Castro, 2001). Pensando-se a variabilidade como efeito-escola, a inclusão de variáveis relevantes forneceria um modelo explicativo das reais diferenças entre pessoas e instituições (Ferrão & Fernandes, 2003).

No que concerne à investigação de fatores associados à eficácia escolar, cabe destaque aos estudos como de Soares (2004, 2005) anteriormente apresentados, em que é enfatizada a importância de controlar fatores relacionados ao próprio estudante e à família (efeito A) e àqueles relacionados estrutura escolar (efeito B). Nesse estudo, será enfatizada a importância de que as avaliações sejam válidas e precisas para inferir com qualidade sobre a eficácia escolar e que possam, para isso, discernir em suas medidas tanto os fatores relacionados à estrutura escolar (contexto) quanto variáveis extra-escolares (como cognitivos e socioeconômicos).

Ao realizar o levantamento bibliográfico em dissertações, teses e artigos científico constata-se que muitas são as variáveis extra-escolares apontadas nas investigações relacionadas à eficácia escolar, e tais escolhas reportam especificidades dos vários contextos escolares e culturais em que são desenvolvidas tais pesquisas. Uma das literaturas de destaque nessa temática foi organizada por Brooke e Soares (2008) a fim de difundir textos teóricos e de pesquisa, históricos e contemporâneos, provindos do cenário

internacional; assim como detalhar alguns avanços nessa temática em pesquisas na América Latina e, particularmente, no Brasil.

Nessa perspectiva, a primeira seção do livro trata de estudos que se iniciaram com o Relatório Coleman produzido em meados em 1960 e que mantiveram a mesma ênfase em ressaltar que a escola tem uma influência comparativamente pequena na explicação do desempenho do aluno. Coleman (2008) aponta que as instituições de ensino são muito similares no efeito que exercem no desempenho de seus estudantes quando levado em consideração o efeito socioeconômico. E, que, é possível entender o desempenho como mais decorrente de fatores combinados das características de instalação da escola e diferenças raciais. Além disso, a qualidade dos professores tem uma associação progressivamente maior com o avançar da escolaridade. Por fim, esse autor conclui que as variâncias relacionadas ao perfil cognitivo e a atitude dos estudantes frente a seu estudo e aprendizagem têm uma relação mais forte com o desempenho do que todos os outros fatores escolares juntos. Tais resultados enfatizavam que haveria maior diferença entre os alunos de uma mesma escola do que entre alunos de escolas diferentes. Porém, cabe destacar que tais conclusões foram tomadas a partir de estudos transversais e não longitudinais.

No contexto dos Estados Unidos, estudos realizados por Jencks (2008) demonstram que o sucesso escolar poderia acontecer por meio de alterações no tamanho da turma, número de horas letivas, organização da escola e qualificação dos professores, além dos aspectos relacionados à motivação dos estudantes. De acordo com esse autor, somente 2% da variação no aproveitamento educacional dos alunos se deve as diferenças qualitativas entre as escolas.

Bowles e Gintis (2008) trazem uma argumentação mais teórica a respeito da incapacidade da escola de alterar a estratificação social americana. Ressaltam a visão da função capitalista da escola em promover uma associação entre produtos e trabalho, e que essa função impossibilita a escola o cumprimento de seus papéis sociais. Nesse capítulo também são feitas distintas revisões de pesquisa a fim confirmar a relação direta entre desigualdade econômica e educacional.

Outros autores como Flouri, Buchanan e Bream (2002), Opdenakker e Van Damme (2000) e Opdenakker e Van Damme (2005), têm indicado que as avaliações escolares deixam de enfatizar a importância de outros aspectos além dos cognitivos em suas medidas e que poderiam derivar em estudos diferenciados nesse sentido. Tais autores ressaltam que, por exemplo, as características do ambiente familiar como o envolvimento dos pais nos estudos dos filhos são variáveis importantes, e podem ser consideradas como um bom preditor de competências acadêmicas.

Porém, Brooke e Soares (2008) destacam que, decorrente dessas visões, se alastrou a tendência em enfatizar em vários estudos que a escola não faz a diferença. E, outros estudos surgiram com a intenção de revisar os métodos, instrumentos e variáveis utilizadas, assim como para reafirmar a capacidade e papel da escola. Madaus, Airasian e Kellagan (2008) argumentam que são as variáveis do processo (como interação entre os alunos) que chegam ao cerne da experiência escolar e ao ponto em que as escolas realmente exercem impacto sobre o ensino e a aprendizagem.

Rutter e cols. (2008a, 2008b) centralizaram-se na discussão de novas metodologias de pesquisa e principalmente na necessidade de saber sobre a situação inicial (anterior) dos alunos para avaliar a aprendizagem e outros ganhos devidos à escola. Também avaliaram estudantes de escolas secundárias de Londres (no Brasil correspondente ao final do Ensino

Fundamental até Ensino Médio) e apontaram fatores importantes de diferenciação entre as instituições, como o ambiente favorável à obtenção de resultados positivos. Esses autores, em suma, destacaram existem variações no comportamento e desempenho dos estudantes entre as escolas, de forma que as escolas que produziam nos estudantes bons comportamentos também produziam maior sucesso escolar, mas que essas diferenças produzidas pelas instituições tendem a ser estáveis ao longo do tempo.

Na Inglaterra, os autores Mortimore e cols. (2008a, 2008b) acompanharam em um estudo longitudinal a trajetória acadêmica de dos mil estudantes dos 7 aos 11 anos de 50 diferentes escolas do país durante quatro anos. As características anteriores dos estudantes foram controladas para evitar a contaminação dos resultados frente à realidade do impacto escolar. Foram realizadas avaliações do ambiente da sala de aula e da escola, levantamento de variáveis relacionadas à família e as condições socioeconômicas, além da utilização de testes cognitivos padronizados de leitura e matemática, e medidas de comportamento entre os estudantes. Concluiu-se, que há a necessidade das escolas serem tratadas diferencialmente, pois houveram diferenças acentuadas entre os estudantes no momento de admissão escolar quando as escolas são comparadas em relação ao progresso e ao desenvolvimento de seus estudantes. Esses autores, porém ressaltam a dificuldade nas pesquisas educacionais em geral para se conseguir medidas de entrada válidas para controlar as diferenças pré-existentes. Apesar desses resultados sobre a influência do perfil anterior dos estudantes, os autores também ressaltam o impacto que as diferentes escolas têm no progresso e desenvolvimento dos seus estudantes (cerca de 25% da variância). Apesar da influência das instituições cursadas serem menores quando comparadas às demais medidas, tais resultados encontrados por Mortimore e cols. (2008a, 2008b) se

mostram superiores a grande parte dos estudos a respeito dos efeitos da escola, que apontam geralmente 2 a 3% da variância.

Lee (2008) apresenta a metodologia dos modelos hierárquicos e a importância da utilização desses em pesquisas educacionais que almejam avaliar o tamanho da agregação escolar. Entre as contribuições apontadas, enfatiza que os modelos hierárquicos permitem a decomposição da relação observada entre o desempenho e o nível socioeconômico em dois componentes, sendo um associado ao aluno e o outro à escola. Também destaca que os modelos hierárquicos permitem estudar situações em que a relação das variáveis explicativas com o desempenho é diferente entre as escolas.

Por fim, Torrecilla (2008) traz um panorama da pesquisa ibero-americana sobre a eficácia escolar e retrata que a maioria dos pesquisadores educacionais da América Latina tem dificuldades em aceitar a legitimidade das medidas de eficácia escolar que apontam efeitos mínimos das instituições na trajetória escolar dos estudantes. Apesar de apontar a necessidade de mudanças nas medidas utilizadas, esse autor destaca a importância da continuidade dessas pesquisas no Brasil e ao longo da América Latina, por acreditar que as mesmas permeiam o debate educacional atual. Segundo ele, esse debate refere-se à qualidade da educação e a identificação de políticas que induzem uma educação de qualidade. Mais especificamente, o autor ressalta que a temática da eficácia escolar tornou-se uma necessidade ao observar que, apesar das dificuldades, as pesquisas com a eficácia escolar já estão presentes em vários países nas pesquisas educacionais.

A fim de corroborar tais apontamentos, Alves e Franco (2008) apresentaram uma síntese das pesquisas realizadas sobre eficácia escolar no Brasil. Segundo esses autores, tais estudos demonstram o interesse de economistas, autoridades e da sociedade em geral na busca por orientações e sugestões de políticas escolares e não-escolares visando uma

melhoria da qualidade de ensino dos estudantes brasileiros. Diante disso, postulam que o papel do pesquisador seria de identificar as características dos estudantes que entram nas escolas, a fim de demonstrar que os resultados diferentes entre uma escola e outra se dão tanto pela diferença entre os alunos quanto pelas diferenças entre as escolas.

Ambos os textos (Alves & Franco, 2008; Torrecilla, 2008) apontam que essa temática despontou no Brasil especialmente na década de 1990, a partir dos dados de desempenho produzidos pelo SAEB. De maneira geral, os autores apontam que os trabalhos realizados no Brasil confirmam os resultados internacionais em que o fator socioeconômico é preponderante na diferenciação dos estudantes das várias instituições. Também destacam que a variância associada ao pertencimento escolar, depois de controlado o nível socioeconômico se mostra um pouco maior do que os valores observados internacionalmente (chegando a 25% em alguns estudos por eles levantados). Em relação aos fatores da eficácia escolar assinalados, juntamente com os demais apontados em outros países, destacam no cenário brasileiro a importância de elementos mais associados a esse contexto, como os recursos humanos e materiais, a qualidade dos docentes e de suas condições para desempenhar o trabalho.

Em geral, o livro organizado por Brooke e Soares (2008) trata de pesquisas realizadas no ensino infantil e fundamental, principal na primeira etapa do Ensino Fundamental, especialmente na Inglaterra e Estados Unidos. Também, do ponto de vista do delineamento, apontam que muitos estudos foram e ainda são baseados em pesquisas transversais e não longitudinais. Nesse sentido, Rutter e cols. (2008a) e Willms (2008) destacam que muitas dos estudos sobre eficácia escolar partem de pesquisas transversais, em que vários ajustes estatísticos precisam ser feitos para estimar o efeito o tamanho do efeito da escola, pois faltam informações sobre os alunos no momento de entrada da escola.

Esses autores então destacam que as medidas de efeito da escola só podem ser determinadas a partir de estudos longitudinais e que apesar de haver maneiras de se lidar estatisticamente esse problema, nenhuma alternativa é realmente tão satisfatória e precisa quanto o uso de medidas de desempenho de um mesmo aluno, antes e depois de ingressar na escola. Por fim, Oliveira, Franco e Soares (2007) também discutem a importância de que os estudos que pretendem avaliar a eficácia escolar e promover a equidade escolar partam da coleta de informações de pesquisas longitudinais, e que somente tal formato de pesquisa oferece dados para estudar a aprendizagem, e a responsabilidade do professor e da escola no processo educacional.

No que se refere aos estudos de análise da qualidade da educação brasileira, verifica-se a criação de diversas instâncias de avaliação, abrangendo diferentes níveis e sistemas de ensino. Ao realizar o levantamento bibliográfico porém despontam as produções vinculadas à análise dos indicadores do SAEB (Sistema de Avaliação da Educação Básica), que se destina a avaliar a aprendizagem dos alunos de 4^a e 8^a séries do Ensino Fundamental e a 3^a série do Ensino Médio por meio de uma metodologia transversal. Esse sistema iniciou-se em 1990 e a partir de 1995 foi introduzido o uso da TRI para a construção de instrumentos, a atribuição dos escores e a análise, de forma a viabilizar a comparação dos resultados. Os últimos dados do SAEB são referentes a 2005.

Devida a extensa literatura sobre o SAEB e as medidas dele equalizadas, e por esse não ser o foco do presente estudo, serão destacadas aqui algumas das pesquisas consideradas mais relevantes, bem como apresentadas as ênfases investigadas e suas principais conclusões. A fim de sintetizar as informações, serão apontados os objetivos do estudo, quais as variáveis intra e/ou extraescolares enfatizadas, o tipo de delineamento (quando pesquisas equalizadas com o SAEB), e principais resultados.

A questão básica ao levantar alguns dos estudos com o SAEB é entender o quanto da variância entre os estudantes é atribuível a escola e a outros fatores (e quais seriam esses) nesses trabalhos. E, quais delineamentos são utilizados para se chegar a essas respostas (transversal e longitudinal), considerando diferentes tipos de variáveis extras e intraescolares. A partir disso, pretende-se comparar os achados com os estudos e propostas voltados para o Ensino Superior, foco do presente trabalho.

Primeiramente, cabe destacar algumas limitações do exame apontadas por Soares e Collares (2007). Segundo esses autores apesar desse sistema ter sido planejado para o monitoramento do ensino, não coleta todas as informações que permitiriam medir de forma adequada os construtos explicativos do desempenho cognitivo dos estudantes, isto é, capital econômico, capital cultural e capital social das famílias dos alunos. Além disso, como os questionários são preenchidos diretamente pelos estudantes, faltam informações importantes, deixadas em branco, provavelmente por desconhecimento dos mesmos sobre os pontos levantados. Também, apontam que, como os dados provêm de respostas dos estudantes sobre suas respectivas famílias, têm-se apenas a visão desses sobre o valor da característica familiar, e não sua medida exata.

Ferrão e cols. (2001) ao investigarem os componentes associados ao efeito escola no Brasil com os dados do SAEB apontaram algumas variáveis específicas e mais pertinentes à realidade brasileira como defasagem da idade com a série escolar, raça dos estudantes, ambiente educacional, recursos didáticos à disposição dos professores e sistema de organização do ensino. Porém, Ferrão e Beltrão (2001) assinalaram que as regiões do Brasil em que a proporção de estudantes com idades acima da adequada para sua série, o desempenho desses estudantes é inferior aos com idade adequada, mas seu efeito é

aleatório, sugerindo que é maior o efeito dessa variável em algumas escolas do que em outras.

Ferrão e Fernandes (2003) fizeram uma revisão dos trabalhos de efeito-escola, a maior parte deles executada com o SAEB e concluíram que os estudos realizados no Brasil apontam para uma tendência da reprodução de desigualdades sociais, em que os desempenhos inferiores se mostraram associados às questões sociais menos privilegiadas. Esses achados também foram corroborados por Ferrão (2003) e Ferrão, Beltrão e Santos (2002). Entre os fatores atrelados ao desempenho escolar foram destacados pelos autores, por exemplo, os aspectos familiares e de conhecimento prévio dos estudantes.

Soares e colaboradores (2002) identificaram, analisando os dados do SAEB do ano de 2001, que estudantes cujos professores cumpriram a maior parte do programa da disciplina apresentaram classificações mais elevadas, com efeito, maior em Matemática, estando associada ainda a menos repetência nas salas. De forma semelhante, César e Soares (2001) analisaram os dados do SAEB de 1999 referentes aos estudantes da 8ª série do Ensino Fundamental que fizeram a prova de Matemática (17.890 estudantes) por meio do modelo de regressão hierárquica, e foram ajustadas variáveis explicativas nos níveis do estudante, escola e estado. Esses autores discutem que a maior parte da variabilidade entre instituições está associada ao perfil socioeconômico dos estudantes e a interação entre os colegas de classe (como atributos ou variáveis de contexto que influenciam no processo de aprendizagem).

Soares e cols. (2004) analisaram os dados produzidos na edição 2003 do SAEB e estimam que o efeito das escolas brasileiras (com o controle da situação socioeconômica dos alunos) varia em média entre 14% e 22%, dependendo da série e da disciplina (matemática ou língua portuguesa). Os autores ressaltam que esses valores são bem

superiores aos observados na maioria dos países desenvolvidos, por exemplo, os que participaram do Programa Internacional de Avaliação dos Alunos (PISA), que giram em torno de 10%.

Soares e Collares (2006) analisaram os dados do SAEB dos alunos de oitava série das escolas públicas, submetidos ao teste de matemática, que perfazem um total de 30.354 alunos divididos em 1.692 escolas de todas as regiões brasileiras. Os resultados das análises de regressão estrutural mostraram que os recursos econômicos da família têm impacto direto apenas sobre os fatores atraso escolar e recursos culturais disponíveis na residência do estudante. E que, embora os recursos econômicos não tenham apresentado efeito direto na proficiência dos estudantes, esses recursos têm efeito indireto na proficiência através das variáveis atraso escolar e recursos culturais, e assim na atitude em relação à escola. Dentre todos os construtos considerados, aqueles que apresentam a associação mais forte entre si são os fatores cultural e econômico da família. Os autores concluem que a análise de dados de pesquisas educacionais deve considerar a influência familiar não através de um único índice síntese de posição social, mas de dois fatores, um medindo os recursos econômicos e outro os recursos culturais da família dos estudantes.

Rauter, Laros e Jesus (2007) analisaram a estrutura fatorial dos dados do SAEB 2001 e, dentre outras questões, verificaram limitações referentes à abrangência do conteúdo da prova. Andrade e Laros (2007) investigaram o desempenho escolar no SAEB de 2001 utilizando análise multinível, especificamente os dados referentes aos estudantes do 3º ano do Ensino Médio. Depois de controlar o efeito das variáveis de seletividade e composição escolar verificou-se que 17% da variância no desempenho escolar podem ser atribuídos ao nível da escola. Verificou-se que o atraso escolar e comparação do estudante com os colegas foram as variáveis individuais com maior poder explicativo. Quanto à escola, a

maior explicação se deu por medidas de recursos culturais agregados e atraso escolar agregado.

A fim de avaliar o efeito-escola a partir de medidas socioeconômicas mais detalhadas, Soares e Andrade (2006) estudaram do efeito de escolas de ensino fundamental e médio de Belo Horizonte/MG, partindo de diferentes medidas dos mesmos estudantes. Os dados vieram do Sistema Mineiro de Avaliação Educacional – SIMAVE, aplicado em 2002 nas escolas estaduais e municipais de Belo Horizonte, e dos vestibulares da Universidade Federal de Minas Gerais (UFMG), referentes aos anos de 2002, 2003 e 2004. O SIMAVE almeja medir o desempenho acadêmico dos alunos do ensino básico de Minas Gerais na escala do SAEB. O indicador de qualidade da escola usado no artigo é o efeito da escola obtido após o ajuste aos dados de um modelo hierárquico de regressão, em que a variável resposta é a proficiência dos alunos. A análise da qualidade das escolas, controlando a influência do nível socioeconômico no desempenho dos alunos, mostrou que algumas escolas (públicas e privadas), pelas suas políticas e práticas pedagógicas conseguem fazer diferença no desempenho de seus alunos mesmo quando eles são socioeconomicamente desfavorecidos. Apesar de apontar esses resultados, os autores não destacam qual a proporção que resta para escola quando controladas as variáveis socioeconômicas. E, da mesma forma que ocorre nos estudos com o SAEB, as interpretações são provenientes de dados transversais.

Diferentemente dos estudos transversais até o momento apresentados, Alves e Soares (2007a) apresentaram os resultados de uma pesquisa longitudinal para medir o efeito das escolas de sete estabelecimentos de ensino públicos em uma região de Belo Horizonte. Os testes e questionários foram respondidos pelos 1194 alunos, seus familiares e profissionais das escolas. Aos alunos, a informação de medida de aprendizagem foi tomada

pelo Banco de Itens do Grupo de Avaliação e Medidas Educacionais com base nos dados do SAEB e nos dados do Projeto Nova Escola, elaborados pelos professores das escolas. Foram realizadas três coletas de dados: duas no primeiro ano da pesquisa, quando os alunos estavam na 5ª série; e a última no final do segundo ano letivo, com os alunos na 6ª série. Em cada uma das aplicações, todos os alunos responderam a testes de língua portuguesa (com ênfase em leitura e interpretação de texto) e de matemática e a um questionário contextual, para levantar informações sobre as características demográficas, socioeconômicas e culturais dos alunos e de suas respectivas famílias, itens sobre o percurso escolar e hábitos de estudo. A técnica de análise empregada foi a TRI, estimando os escores longitudinais dos alunos e seus níveis socioeconômicos e os dados foram equalizados nas provas escolares, respondidas pelos alunos. A análise do efeito das escolas nos dados longitudinais de dois anos de acompanhamento pela modelagem multinível evidenciou que a incorporação do nível socioeconômico foi o mais forte preditor do desempenho escolar, de forma que independente do contexto, os alunos com melhor nível socioeconômico apresentaram maior probabilidade de beneficiarem-se de suas vantagens extra-escolares e apresentarem resultados melhores; sendo que as melhores escolas mostraram uma tendência de estabilidade no período final do estudo, reproduzindo os bons resultados iniciais. Os autores optaram por apresentar os resultados dos efeitos das escolas somente em termos de positivo ou negativo, e se muito forte, forte, moderado ou não significativo e justificaram essa opção por não se tratar de um estudo experimental. Partindo do mesmo banco de dados e da metodologia multinível para verificação do efeito-escola, Alves e Soares (2008) mostraram que os estudantes com nível inicial mais baixo, em média, progrediram mais que aqueles com nível inicial mais alto, embora destaquem que essas diferenças de nível persistam no final do estudo. De uma forma geral, as escolas se tornaram muito mais

parecidas no final do estudo, embora ainda com diferenças significativas. As escolas que tinham médias iniciais mais altas cresceram menos que as escolas com médias iniciais mais baixas, o que para os autores deixa evidente que o efeito das escolas no aprendizado dos estudantes pode ser subestimado com dados transversais. Segundo eles, os atributos socioeconômicos têm impacto significativo na explicação de diferenças no nível inicial, mas apenas algumas têm impacto no aprendizado dos alunos, dependendo da disciplina e da série.

De forma semelhante, Alves e Soares (2007b) realizaram uma pesquisa longitudinal com três coletas de dados, durante os anos de 2003 e 2004, de sete escolas públicas relativamente homogêneas do ponto de vista socioeconômico. A amostra inclui alunos de 28 turmas de sete escolas, perfazendo um total de 740 alunos que fizeram o teste de língua portuguesa e 749 que fizeram o de matemática. O efeito-escola foi mensurado através da identificação da parte do progresso dos alunos que pode ser atribuída às escolas. Neste artigo, analisaram os dados obtidos nas duas primeiras aplicações, que ocorreram durante o primeiro ano letivo do estudo, quando os alunos freqüentavam a 5ª série ou ciclo equivalente. Em cada onda, todos os alunos de todas as turmas respondem a testes de língua portuguesa (com ênfase na leitura e compreensão de textos), de matemática e a um questionário contextual para levantamento de informações sobre as características demográficas, socioeconômicas e culturais dos alunos e de suas respectivas famílias, além de itens sobre o percurso escolar e hábitos de estudo. Os resultados revelam que o efeito-escola se estrutura pela forma como as escolas organizam os alunos em turmas. A variabilidade entre as turmas para os ganhos em matemática, que já era pequena no modelo nulo, ficou praticamente a mesma (de 6% para 5%). Mas em língua portuguesa, o controle pelo contexto das turmas reduziu a variabilidade entre as turmas de 16% para 12%. Os

autores concluem que a formação das turmas por nível de habilidade impacta não apenas o nível de desempenho dos alunos, mas também as diferenças nas taxas de progresso observadas durante o estudo.

Cabe destacar que os estudos de Alves e Soares anteriormente apresentados permitem o questionamento sobre as formas de acompanhamento escolar e de diferenciação de instituições, como a proposta do Projeto GERES, estudo Longitudinal da Geração Escolar. Além disso, trata-se de estudos que antecederam o projeto GERES. O GERES é estudo longitudinal de alunos de ensino fundamental em cinco grandes centros urbanos brasileiros, a saber: Belo Horizonte, Campinas, Salvador, Campo Grande e Rio de Janeiro e seus principais objetivos envolvem a identificação das características escolares e práticas educativas que maximizam o aprendizado e minimizam a desigualdade na distribuição social do aprendizado. A mensuração da aprendizagem dos alunos foi programada para 5 etapas diferentes durante quatro anos e iniciou em março de 2005 com a análise do nível de proficiência com que os alunos entraram na escola. Além das investigações cognitivas de português (leitura) e matemática, foram aplicados questionários aos alunos, pais, professores e diretores. E, alguns trabalhos foram localizados decorrentes dos dados do GERES, como se segue.

A fim de discutir formas de medir e produzir medidas de nível socioeconômico (NSE) em pesquisas longitudinais, Alves e Soares (2009) utilizaram os instrumentos de avaliação contextual da pesquisa GERES. O resultado obtido mostrou que o nível socioeconômico calculado a partir das informações sobre a escolaridade, ocupação e renda discrimina melhor as famílias ao longo da escala de NSE. O artigo de Oliveira, Franco e Soares (2007) concentrou-se em apresentar a metodologia utilizada no Projeto GERES e avaliar a qualidade dos itens por meio da TRI, além da análise pedagógica de cada um dos

itens. Semelhantemente, Stevão (2008) e Ubriaco (2009) utilizaram-se da TRI para a interpretação de escalas de medida da competência matemática do GERES. A importância dos resultados e implicações das medidas proporcionadas por esse projeto foi destacada por Alves e Soares (2008). Apesar de esse projeto ter como um dos propósitos a geração de estudos sobre a eficácia escolar, não foram encontradas pesquisas que assim dispusessem dessas informações.

Ao analisar as pesquisas encontradas, verifica-se que grande parte dos estudos está baseada em medidas transversais e avaliada no nível do aluno. Os objetivos centram-se na discussão de quais medidas devem ser utilizadas para avaliar e controlar o nível socioeconômico dos estudantes, resultados de análises de regressão multinível para mensuração do efeito-escola e utilização da Teoria de Resposta ao Item para as análises psicométricas e interpretação dos resultados. Em geral, verificou-se um consenso de que o desempenho acadêmico de estudantes de ensino fundamental de escolas públicas é mais explicado por fatores extra-escolares do que intra-escolares. Com efeito, os resultados escolares mostram-se mais associados às características socioeconômicas e culturais dos alunos, e ao perfil anterior a entrada do aluno, sendo o efeito das escolas (após o controle) de no máximo 25%.

Tais estudos, porém devem ser avaliados com cautela para não conduzir a suposição de que a escola pouco afetaria o desempenho dos estudantes, e portanto, levar ao entendimento de que não vale à pena investir em formas de organização escolar. Nesse aspecto, autores como Rutter (1979) discorrem sobre a variação de desempenho dos estudantes explicada por fatores internos à escola, como o aumento do desempenho quando os estudantes frequentam escolas com características associadas a um melhor desempenho e a importância do relacionamento com os colegas. Nesse sentido, é fundamental investigar

quais fatores realmente estão associados à eficácia escolar após controlar outras variáveis confundidoras (Soares, 2004).

Outro instrumento de destaque é o Exame Nacional do Ensino Médio (ENEM) aplicado aos estudantes concluintes desse nível de ensino. Da mesma forma que realizado com os estudos do SAEB, serão destacados alguns dos estudos associados aos resultados do ENEM. Mildner e Silva (2002), e Vianna (2003) de forma quantitativa e qualitativa, trataram sobre a escassez de estudos de validade desse exame e a necessidade de melhor clareza nas interpretações propostas. Alguns estudos encontrados tiveram como objetivo a análise do modelo conceitual de habilidades e competências que é subjacente ao ENEM à luz de modelos contemporâneos da inteligência (Gomes, 2005; Primi & cols., 2001). Outros estudos abordaram a respeito da construção de indicadores da qualidade de ensino, variáveis associadas e a repercussão do exame no contexto escolar e social (Camargo, 2007; Castro, 2000; Cerri, 2004; Colombi, 2004; Maggio, 2006; Machado, 2003; Oliveira, 2006; Riani, 2005; Santos Neto, 2006; Sousa, 2003; Souza & Oliveira, 2003; Zanchet, 2003; Zironi, 2007). Em geral, as análises relacionadas ao ENEM estão mais voltadas ao entendimento do construto avaliado pelas provas e ao impacto desse exame no acesso ao Ensino Superior, não sendo o objetivo do exame a avaliação do efeito-escola.

Convém investigar como as questões levantadas nos estudos do Ensino Fundamental, cujo foco é a mensuração do efeito-escola, se assemelham aos apontamentos e investigações realizadas no Ensino Superior. Entende-se que para a avaliação da trajetória dos estudantes do Ensino Superior também se faz necessária a existência de medidas válidas de controle inicial do conhecimento dos estudantes, bem como das variáveis socioeconômicas, além de interpretações baseadas em medidas longitudinais. O levantamento bibliográfico realizado apontou a existência de estudos que discutem, por

exemplo, os conceitos e as propostas de avaliações em larga escala como o ENADE (Barreyro, 2004; Bittencourt, Viali, Casartelli, & Rodrigues, 2008; Limana & Brito, 2005; Verhine, Dantas & Soares, 2006; Vianna, 2003); questões associadas aos perfis de cursos avaliados pelo ENADE (Brito, 2007; Lopes, Vendramini & Silva, 2010; Silva, Vendramini & Lopes, 2010); além de questões psicométricas e limitações do exame para a avaliação da qualidade dos cursos das instituições (Burlamaqui, 2008; Nogueira, 2008; Oliveira, 2006; Paiva, 2008; Primi, 2006b, Primi & cols., 2009a, 2009b; Primi & cols., no prelo; Primi, Hutz & Silva, no prelo; Schwartzman, 2010; Vendramini, 2005; Vendramini, 2007; Vendramini, Silva, Dias & Nogueira, 2007).

Porém, poucos estudos empíricos foram encontrados com a ênfase de investigar o efeito-escola no Ensino Superior, cabendo destaque ao trabalho desenvolvido por Soares, Ribeiro e Castro (2001). Esses autores analisaram o valor agregado das IES de Minas Gerais dos cursos de Direito, Administração e Engenharia Civil, mediante a junção das bases de dados referentes contendo as notas dos estudantes na primeira etapa dos vestibulares da UFMG entre 1992 e 1996 e o banco de dados do INEP, com as notas dos estudantes do Provão de 1996 a 1999. Os pesquisadores se utilizaram de análises de regressão hierárquica a fim de incluir variáveis explicativas do desempenho prévio dos estudantes, nível socioeconômico, além de fatores individuais e de contexto (como efeito dos pares na aprendizagem). Considerando o desempenho prévio, sexo e nível socioeconômico, o efeito das instituições no desempenho para os 3 cursos é de 6%, como maior influência da medida de desempenho inicial nesse modelo. Quando controlado também o efeito entre os pares, a variância abaixa para 2% no curso de Engenharia e 3% para o Direito. Concluiu-se que, devido à grande variabilidade no perfil de entrada dos estudantes e da influência significativa das características dos estudantes sobre seu

aprendizado, é fundamental que se incorporem medidas das condições iniciais dos estudantes, a fim de que seja possível observar com maior precisão a efetividade das IES's no Brasil. Uma das sugestões feitas seria de que as informações do ENEM pudessem servir como uma medida nacional do desempenho prévio (ou de entrada) no Ensino Superior, sendo para isso, o seu uso como uma medida generalizada nos processos de admissão das instituições. Outro questionamento levantado nesse estudo, é que mesmo com o uso de medidas longitudinais, verifica-se que existe pouca mudança devido às instituições.

Cabe destacar que a carência de estudos empíricos e com os devidos controles metodológicos e estatísticos sobre as variáveis extra-acadêmicas mais relevantes nas pesquisas de efeito-escola no Ensino Superior é um ponto a ser discutido e melhor investigado. Igualmente, Brooke e Soares (2008) destacam que no Brasil a ausência de mensuração do conhecimento prévio do aluno leva os pesquisadores a assumir, nos modelos de análise, a forte correlação entre 'conhecimento prévio e nível socioeconômico e/ou atraso escolar' dos alunos no momento de entrada na escola. E que embora esses apontamentos sejam razoáveis para analisar o nível de desempenho escolar, tal metodologia não permite afirmações válidas e precisas sobre o efeito das escolas no processo de aprendizagem.

Além das questões relacionadas a quantidade de variabilidade associada realmente a qualidade dos cursos das instituições e a verificação da importância dos fatores individuais na variação observada entre as proficiências dos estudantes (Cesar e Soares, 2001), o capítulo seguinte irá tratar sobre o entendimento do significado do desempenho proposto nessas medidas avaliativas em termos de construto e, assim, de que maneira avaliações, como o ENADE, podem ser entendidas a luz das teorias de inteligência.

3. SUBSÍDIOS PSICOLÓGICOS PARA O ENTENDIMENTO DAS MEDIDAS DE APRENDIZAGEM

Uma das questões essenciais nos contextos educacionais é como verificar se os estudantes adquiriram o conhecimento transmitido e quais as habilidades e competências necessárias. Nesse âmbito, Urbina (2007) ressalta que o primeiro uso da testagem no campo da educação ocorreu na Idade Média, com o surgimento das primeiras universidades européias no século XIII. E, que, por volta dessa época, o nível universitário passou a ser usado como forma de caracterizar a qualificação para ensinar, e os exames formais aplicados a fim de dar oportunidades aos candidatos de demonstrar sua competência. No final do século XIX, tanto na Europa como nos Estados Unidos, estes exames já estavam bem estabelecidos como método para determinar quem deveria receber um diploma universitário e quem estava capacitado a exercer profissões liberais, como medicina ou direito. Porém, as testagens educacionais, geralmente baseadas na avaliação dos diferentes níveis de desempenho, permitem alguns questionamentos centrais, tais como: quais construtos são medidos por esses testes? Como eles são concebidos nas teorias da inteligência? Considera-se que a explanação desses temas auxiliam na compreensão do construto medido pelo ENADE.

Nessa temática, convém citar os destaques feitos no texto de Alves e Soares (2008) no que diz respeito aos estudos sobre desempenho escolar (*achievement*) e aprendizado (*learning*). De acordo com esses autores, na área educacional, geralmente os estudos sobre a primeira noção estão voltados ao estabelecimento de um determinado nível de desempenho, enquanto que o aprendizado ou progresso dos alunos, diz respeito à aquisição de conhecimentos, de habilidades, de crescimento intelectual ou físico ou de mudanças

(nem sempre positivas) que ocorrem na trajetória escolar. Assim, para esses autores, o entendimento da dimensão aprendizado é central para a educação, porque reflete os processos que ocorrem com os estudantes durante o tempo em que eles estão frequentando determina instituição de ensino.

A partir de uma visão interacionista da inteligência, a aprendizagem e, conseqüentemente, a experiência, se constituem como um aspecto necessário para a aquisição de novos conhecimentos. Por essa razão, o entendimento de como se dá aprendizagem se encontra inevitavelmente unido ao estudo da inteligência. Também é preciso considerar que a aprendizagem igualmente resulta em abertura de possibilidades decorrentes das novidades provindas do meio (Dongo, 2008).

De forma complementar, pode-se entender que para que a aprendizagem aconteça devem estar envolvidos três elementos, o aluno, o processo de aprendizagem e a situação ou contexto de aprendizagem, cabendo então avaliar todos esses elementos ao intentar determinar as influências no desempenho (Falcão, 1985). Assim, o termo aprendizagem implica em aquisição de conhecimento e mudança de comportamento. Também é preciso considerar que a aprendizagem refere-se a qualquer mudança relativamente permanente nos comportamentos e/ou sentimentos de uma pessoa resultante da experiência (Sternberg, 1997).

De forma geral, pode-se conceber que a aprendizagem aumenta a capacidade de adaptação do organismo e refere-se a um aspecto específico da inteligência humana (Primi, 2008). Segundo esse mesmo autor, a aprendizagem é o processo de aquisição de informações ou à organização de informações na memória de longo prazo (procedimental e declarativa) de tal forma que possam ser utilizadas no futuro. Por conseguinte, como será mais amplamente detalhado posteriormente, algo essencial para que o estudante tenha êxito

no processo de aprendizagem se refere a sua habilidade ou potencial para a mesma e a aplicação dos seus conhecimentos no cotidiano escolar. A explanação sobre esse assunto é essencial visto que as diretrizes dos sistemas de avaliação educacionais partem de um conjunto de habilidades consideradas essenciais para o aprendizado.

Habilidade e Competência

Existem algumas diferenciações teóricas em relação ao emprego das terminologias habilidade, capacidade e aptidão. Primi e cols. (2001) ressaltam que, na língua portuguesa, possivelmente, o termo habilidade tenha um significado mais restrito daquele que é atribuído ao termo “*abilities*” na língua inglesa e, que, o termo mais adequado para traduzi-lo seria “capacidade”. No entanto, por uma questão de continuidade com a terminologia usada em trabalhos anteriores, o termo habilidade, quando usado, deve ser entendido como sinônimo de capacidade.

Segundo Carroll (1993a) citado por Primi e cols. (2001), habilidade é a variação individual disponível para a realização de uma classe definida de tarefas e indica a facilidade em lidar com uma determinada informação.. Esse autor concebe que existem algumas habilidades abrangentes que estão no cerne de alguns tipos de tarefas para as quais as pessoas mostram possuir potencial diferenciado. Distingue, ainda, o conceito de habilidade de aptidão afirmando que para esse último é preciso haver maior estabilidade temporal, e, portanto, influência genética.

Alguns autores como Snow (1992) definem aptidão como uma predisposição das pessoas em emitir respostas diferenciais à determinada situação ou classe de situações, significando um estado inicial que irá influenciar o desenvolvimento posterior (dadas determinadas condições). Com relação ao conceito de habilidade, essa última concepção

difere da proposta por Carroll (1993a), tendo em vista que ele usa o termo aptidão, sem se referir ao problema de estabilidade ou não do construto. Salvo as diferenciações quanto à concepção genética, que daria um atributo de estabilidade, a idéia principal é a de que a aptidão envolve também a interação em uma dada situação e, assim, as pessoas possuem aptidões ou habilidades, como referidas neste texto, para se desempenhar em determinadas situações específicas.

De acordo com Primi e cols. (2001) a habilidade é o potencial do indivíduo para realizar determinada tarefa com relativa facilidade, sendo que esse potencial pode ser observado em várias situações, como quando o indivíduo apresenta respostas corretas e conhecimento de determinados conteúdos. Acrescenta-se a isso a importância dada por Brito e cols. (2000) de verificar quais medidas de habilidade estão mais relacionadas ao desempenho em tarefas específicas. Segundo Krutetskii (1976) o estudo das habilidades está intimamente relacionado ao entendimento das diferenças individuais, pois se todos os indivíduos tivessem todas as habilidades desenvolvidas em um mesmo nível, o problema que permeia a avaliação das mesmas deixaria de existir.

No presente estudo o conceito de habilidade será entendido como em Primi e cols. (2001). Esses autores ressaltam que a habilidade, no sentido empregado por Carroll, pressupõe a existência de uma relativa facilidade em lidar com informações e com problemas de uma determinada classe ou conteúdo. É que o investimento deste potencial em experiências de aprendizagem pode conduzir à maestria, que envolve um conhecimento organizado sobre um determinado tema. Nessa perspectiva, a idéia de maestria implica que a realização atingiu um determinado nível a partir do qual se diz que uma competência foi adquirida.

A partir disso, é essencial então definir competência. Para Moreno (2006) o conceito de competência revela um conhecimento em profundidade, reconhecido e que confere o direito de decidir em alguma área. Também para esse autor, uma pessoa competente é aquela que sabe confrontar situações complexas ou resolver problemas utilizando seus conhecimentos e capacidades na execução altamente adequada de uma tarefa.

Bastos (1999) e Miranda (2006) ressaltam que a noção de competência envolveria uma capacidade de ação em diversas situações. Refere-se aos recursos adquiridos ou possuídos que são postos em ação em situações práticas, pressupondo uma mobilização de recursos pessoais. Nesse contexto, ser competente consistiria, basicamente em mobilizar e combinar recursos, integrando diversos conhecimentos para realizar uma atividade de forma eficiente. Miranda (2006) fez uma revisão a respeito dos conceitos de competência e destaca que os autores que comumente estudam as competências em nível individual costumam separá-las em três componentes básicos, a saber, conhecimentos (ter informação, saber o quê e o porquê), habilidades (possuir técnicas, saber como) e atitudes (querer fazer, determinação e identidade), sendo que esses componentes são normalmente designados pela sigla CHA.

Semelhantemente, Parry (1996,1998) destaca que para uma tarefa ser realizada com competência é imprescindível que haja habilidade para executá-la e domínio dos conhecimentos necessários, além de atitude adequada frente à tarefa. E que a competência está sempre associada a uma mobilização de saberes e a capacidade de recorrer ao que se sabe para realizar o que se deseja. Assim sendo, o termo competência está ligado à idéia de desempenho eficiente, ao cumprimento efetivo da ação e à capacidade de realização com êxito de uma determinada tarefa (Brito & cols., 2000).

Ao verificar essas conceituações, é possível pensar que a habilidade indica facilidade em lidar com um tipo de informação e para que se transforme em competência é necessário investimento em experiências de aprendizagem. Considerando a mesma qualidade e montante de experiência, duas pessoas com habilidades diferentes diferirão na facilidade com que se tornarão competentes em determinado tema (Primi, 1998; Primi & cols., 2001).

Assim sendo, no que diz respeito ao presente estudo, entende-se que as habilidades acadêmicas aplicadas a determinadas tarefas, juntamente com a aquisição satisfatória de conhecimentos por meio de experiências adequadas de aprendizagem e do grau de investimento nos estudos, levam a um desempenho altamente satisfatório. Deste modo, um nível elevado de desempenho pode demonstrar a competência do estudante nos conhecimentos específicos da área e, este domínio, pode ser considerado como um dos elementos importantes na formação das competências profissionais (além da importância de outras características individuais, sociais e do contexto). Nesse panorama, os conceitos, competências e habilidades, são fundamentais na compreensão do que o ENADE se propõe a medir, bem como nas interpretações e implicações práticas de seus resultados.

A partir dessas diferenciações, o desempenho pode ser entendido como o produto da interação do processo educacional com as habilidades individuais, e a avaliação um modo de demonstração dos conhecimentos adquiridos por meio do processo de educação formal, que quando altamente satisfatório (maestria), poderia demonstrar eficiência ou competência em determinado conteúdo (Smith & Marsiske, 1994). Na prática, o desempenho é considerado que uma das variáveis observáveis dos construtos latentes habilidade/capacidade entendidos como potenciais para aprendizagem e da competência, entendida como potencial realizado, posto em prática, com eficiência. Verifica-se que

devido ao dinamismo existente entre os conceitos de habilidades e competências, tais significados muitas vezes acabam por ser confundidos, como ocorre em algumas definições propostas pelo próprio ENADE em que essas são muitas vezes colocadas como sinônimos (ver diretrizes no Anexo 2).

A fim de entender quais habilidades estão envolvidas no processo de aprendizagem é preciso recorrer às teorias de inteligência. Dentre as abordagens teóricas da inteligência, o presente estudo explorará mais particularmente a abordagem psicométrica. Um dos alvos privilegiados da psicometria no estudo da inteligência tem sido a identificação das habilidades relacionadas ao desempenho acadêmico (Almeida, 1988; Primi, 2002b).

A teoria das capacidades cognitivas de Cattell-Horn-Carroll (CHC) representa um avanço das abordagens psicométricas e vem sendo usada como referencial para a compreensão da hierarquia das funções cognitivas humanas (Flanagan & Ortiz, 2001). A força acadêmica do modelo foi expressa por autores como Flanagan e McGrew (1998), ao afirmarem que sua robusta rede de evidências de validade tornava-o capaz de dar suporte empírico à elaboração de testes de inteligência. Para Schelini (2006), na última década, o Modelo CHC teve um significativo impacto na revisão de tradicionais testes de inteligência, bem como na elaboração de novos, e tem-se exemplos importantes como o WAIS III e o WISC IV, respectivas revisões da Escala de Inteligência Wechsler para Adultos (WAIS) e da Escala de Inteligência Wechsler para Crianças (WISC).

O modelo CHC compreende a estrutura da inteligência a partir de três camadas hierárquicas dispostas em função da generalidade do conteúdo das habilidades mentais. No primeiro estrato encontra-se um conjunto de aproximadamente setenta fatores específicos ligados ao conjunto de tarefas que são avaliadas pelos testes de um modo geral; no segundo estrato, e relacionando-se aos fatores do primeiro estrato, encontram-se dez fatores amplos

ligados a áreas do funcionamento cognitivo e por último, no terceiro estrato, encontra-se o fator geral da inteligência, indicando a existência de operações cognitivas comuns a todas as atividades mentais dos três estratos (Carroll, 1993b, 1997; McGrew & Flanagan, 1998; McGrew, 2003).

Os dez fatores amplos do modelo CHC, localizados no segundo estrato do modelo, são descritos na teoria seguindo uma ordem hierárquica que indica o nível de associação entre cada fator amplo e o fator geral de inteligência. Assim, seguindo essa hierarquia, Cattell, Horn e Carroll classificaram e nomearam os fatores amplos da seguinte forma: Inteligência Fluida (*Gf*), Inteligência Cristalizada (*Gc*), Conhecimento Quantitativo (*Gq*), Leitura e Escrita (*Grw*), Memória de Curto Prazo (*Gsm*), Processamento Visual (*Gv*), Processamento Auditivo (*Ga*), Capacidade de Armazenamento e Recuperação de Memória de Longo Prazo (*Glr*), Velocidade de Processamento (*Gs*) e Rapidez de Decisão (*Gt*) (McGrew, 2004). Todos esses fatores, mais em especial, *Gf* e *Gc* possuem uma maior associação com o fator geral da inteligência. A primeira habilidade envolve capacidades mentais relacionadas com processos lógicos em situações relativamente novas, enquanto que a segunda abrange capacidades, conhecimentos e conteúdos decorrentes da experiência educacional e exposição cultural necessários para lidar com situações problema (Ackerman, 1996; Phelps, McGrew, Knopik & Ford, 2005; Primi & Almeida, 2002) ou seja, fortemente condicionadas pela aprendizagem e treino (Almeida, 1988; McGrew, 1997; Primi, 2002a, Primi, 2002c). A Figura 2 ilustra a estrutura da inteligência segundo o modelo CHC.

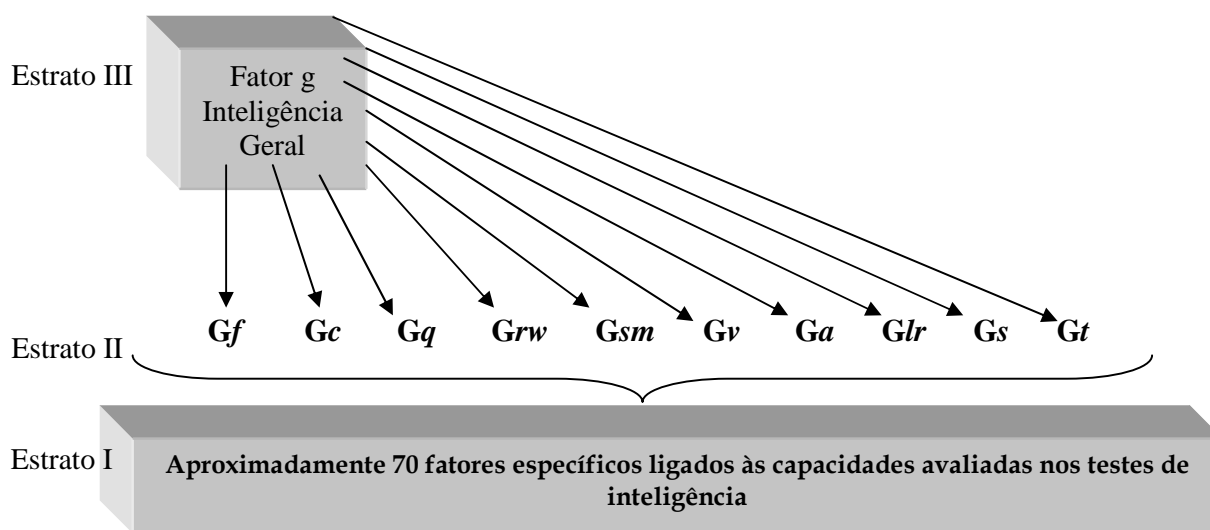


Figura 2. Estrutura da inteligência segundo o modelo de Cattell-Horn-Carroll

Almeida (1994) relata que as habilidades estão organizadas em um *continuum* e *Gc* localiza-se em um dos extremos, correspondendo às habilidades específicas associadas à aprendizagem e aos conhecimentos adquiridos em determinado momento. Assim, em termos conceituais *Gf* é teorizada como uma inteligência mais associada ao raciocínio e estreitamente dependente de fatores biológicos (maturação, envelhecimento, entre outros.). E *Gc* está mais ligada às habilidades desenvolvidas fruto de *Gf*, mas também da experiência, do investimento e das aprendizagens dos indivíduos, ou seja, associada à extensão e profundidade dos conhecimentos acumulados a partir das interações particulares da pessoa com seu ambiente cultural formal e informal (Almeida, 1988; Cattell, 1971; Horn & Cattell, 1966, 1967; McGrew & Flanagan, 1998). Nesta temática, Undheim (1981) destaca que as pessoas poderiam apresentar entre si, um nível igual de *Gf* e diferentes níveis de *Gc*, o que se explicaria pela extensão ou diferentes esforços relacionados ao percurso acadêmico.

Schwiezer e Koch (2002) ressaltam que, existe um multidirecionamento entre inteligência e aprendizagem e, mais especificamente, são propostas fortes relações tanto

entre a inteligência fluída e aprendizagem, quanto entre a inteligência cristalizada e aprendizagem. O pressuposto então é de que o aprendizado controla o acesso da informação como forma de obtenção de conhecimentos básicos e que, os conhecimentos se dão por intermédio da aprendizagem.

A partir da integração apresentada na Figura 2, alguns estudos têm sido desenvolvidos a fim de clarificar, revisar e expandir esta teoria, dentre os quais se destaca a revisão de McGrew (2004, 2005), em que foram incorporadas mais seis capacidades amplas, a saber: Conhecimento em domínios especializados (*Gkn*), Velocidade Psicomotora (*Gps*), Habilidades Psicomotoras (*Gp*), Habilidades Olfativas (*Go*), Habilidades Táteis (*Gh*) e Habilidades Cinestésicas (*Gk*).

Dentre essas capacidades gerais (ou fatores amplos), este estudo tem por objetivo aprofundar-se mais no *Gkn*, que segundo McGrew (2004, 2005) pode ser compreendido como a amplitude e profundidade de conhecimentos adquiridos em domínios especializados (demarcados) e que tipicamente não são representadas por aquelas experiências universais e gerais de indivíduos em uma determinada cultura (*Gc*). Nessa perspectiva, *Gkn* reflete a profundidade dos domínios de conhecimento especializados, desenvolvidos por uma prática sistemática intensiva e treinamento (por um período extenso de tempo) e a manutenção da base do conhecimento por prática regular e esforço incentivado. Portanto, o *Gkn* se configura como uma elaboração das teorias a respeito de *Gc*, podendo assim ser considerada como uma extensão ou especificidade da mesma.

De acordo com Hunt (2000), essa discussão deve ser estendida ao tipo de conhecimento adquirido, em que independentemente da área, o simples fato de possuir um conhecimento amplo sobre alguma área, por mais específica ou global que seja, envolvem a inteligência cristalizada sob manifestações diferenciadas. Apesar de tal afirmação do autor,

essas considerações se mostram próximas da definição de *Gkn*. Convém, no entanto, ressaltar que apesar de considerar a definição de *Gkn* como mais apropriada e específica para entender as propostas de avaliação do Ensino Superior e do desempenho dos estudantes, tal elaboração teórica é bastante recente e, assim, grande parte dos estudos apresentados tratam da relação entre *Gc* e o conhecimento adquirido.

Nesse sentido, pode-se pensar que, ao longo do desenvolvimento cognitivo da infância até ao final da adolescência, se partirá de uma habilidade mais básica e geral (tipo *g* ou *Gf*) para as habilidades cognitivas diferenciadas (tipo *Gc*, *Gkn* e outras aptidões), confiando um papel importante a experiência, a aprendizagem e a motivação dos sujeitos (Almeida, Lemos, Guisande & Primi, 2008). Ackerman (1996a) ressalta tal aspecto ao discorrer que *Gf* tem o seu pico de desenvolvimento no final da adolescência, mantendo-se estabilizada por alguns poucos anos e começando a partir daí a declinar, principalmente nos aspectos relacionados com a velocidade de processamento; *Gc* desenvolve-se com as aprendizagens e a experiência, prolongando o seu desenvolvimento pela idade adulta e permanecendo estável por um período mais longo de tempo que *Gf*, podendo apresentar algum declínio a partir de idades já bastante avançadas (terceira idade).

É relevante tratar dessa temática a fim de investigar se há mudanças desenvolvimentais que poderiam, por conseguinte, ressaltar a necessidade de diferenças na mensuração da inteligência. A maior parte dos trabalhos revisados da literatura que concernem à inteligência de adultos indicou que essa fase é caracterizada por um declínio na velocidade dos processos mentais, assim como da capacidade de raciocínio e algumas medidas de desempenho de memória (Salthouse, 1991). Tal concepção é semelhante à de Cattell (1957) e de Horn e Cattell (1966) de que *Gf* exerce um papel mais determinante na infância, guiando a aquisição de conhecimento cultural e habilidades específicas. Já na fase

adulta, a influência dominante sob o desempenho é mudada gradualmente para *Gc*, sendo que essa última tende a aumentar com o passar da idade ou permanecer constante.

Além disso, conforme os indivíduos se movem no curso do desenvolvimento há um maior contato com ambientes escolares formais. As diferenças no número de anos de educação e o tipo de experiências educacionais que esses anos representam têm sido examinados como um fator importante na compreensão dos efeitos de aprendizagem (Berg, 2000; Parks, Mitchell & Perlmutter, 1986).

Mais especificamente, algumas pesquisas no campo da educação, ciência cognitiva e desenvolvimento, discutem que a maioria dos testes de inteligência se detêm geralmente na análise dos processos (medidas de raciocínio, como inteligência fluida) e menos em conteúdo (domínios específicos de conhecimento), e que especificamente na avaliação de adultos, esse último é mais relevante (Ackerman, 1996a, 1996b; Ackerman, 2005; Ackerman, Beier & Bowen, 2002; Castorina & Faigenbaum, 2002; Gregory, 1994). Essa importante concepção teórica emergiu no final de 1960 e início de 1970, e diz respeito à relevância da consideração das medidas de conhecimento em domínio específico, principalmente com o avançar da idade (Ackerman, 1996a, 1996b; Cronbach, 1957; Hambrick, 2005; Smith & Marsiske, 1994).

Portanto, a partir do conceito de que o indivíduo adquire, amplia e altera sua capacidade ao longo do desenvolvimento, Ackerman (1996a, 1996b) e Ackerman (2000) propõe uma teoria sobre o desenvolvimento cognitivo do adulto, em que refere-se a uma *inteligência como processo* (*gp*) e uma *inteligência como conhecimento* (*gk*). Segundo Ackerman (1996a, 1996b) e Ackerman & Beier (2003), a importância dos dois tipos de inteligência modifica-se ao longo do desenvolvimento, sendo que as habilidades típicas de *Gf* têm o seu cume no final da adolescência e início da vida adulta, começando a declinar

principalmente em termos de velocidade de processamento. Por outro lado, *Gc* se desenvolve gradativamente de maneira mais lenta estabilizando-se no final do período de jovem adulto, com predomínio da importância do domínio de conhecimentos acadêmicos e ocupacionais, e seu declínio se dá numa faixa etária mais avançada (terceira idade). Em termos acadêmicos, esse autor destaca que os *experts* obtêm vantagens sobre os novatos no desempenho em domínios específicos e essas estão relacionadas principalmente aos níveis mais elevados de conhecimento e conhecimento mais integrado.

Diversas linhas de pesquisa convergiram em uma redefinição das concepções a respeito das medidas de inteligência na idade adulta, de forma a considerar mais o conteúdo (conhecimento) nessa fase de desenvolvimento (Ackerman, 1996a, 1996b; Ackerman, 2005; Ackerman, Beier & Bowen, 2002; Castorina & Faigenbaum, 2002; Cronbach, 1957).

De forma congruente, várias pesquisas têm se focado no exame das diferenças entre estudantes novatos (iniciantes, aprendizes) e especialistas (*experts*) em uma determinada área de conhecimento, e apontado que a principal distinção do especialista está na experiência e estrutura de conhecimento, do que em termos de processos intelectuais (tais como, Ackerman, 1996b; Ackerman & Beier, 2003; Ackerman & Beier, 2005; Connell, Sheridam & Gardner, 2003; Sternberg & Grigorenko, 2003). Nessa perspectiva desenvolvimental, pesquisas gerontológicas têm apontado a menor influência dos processos cognitivos e a maior determinação da estrutura do conhecimento no declínio da idade (Berg, 2000; Schooler, 1987; Sternberg, 2000).

Entende-se que a principal contribuição dos estudos de Ackerman centra-se na retomada das medidas de conhecimento e processo para se estudar a inteligência. Ou seja, sua intenção não era de desconsiderar a importância de outras capacidades cognitivas

envolvidas no desempenho em adultos, e sim enfatizar que os testes de conhecimento podem ser considerados como um dos componentes das medidas de inteligência e com maior ênfase nessa etapa de desenvolvimento. Nesse sentido, cabe citar a relevância de pesquisas como a de Ackerman, Beier e Boyle (2005), que têm apontado que capacidades como a memória de trabalho estão altamente relacionadas com o domínio do conhecimento, podendo ser relativamente importante no processo de especialização do conhecimento (em outras palavras, na definição do estudante especialista, *expert* ou mais competente em uma determinada área de conhecimento). Demais pesquisas, como a de Ackerman e Cianciolo (2002), têm investigado outras medidas de inteligência, como habilidade espacial e numérica, além da importância da experiência prática na tarefa, como fatores essenciais para um desempenho bem sucedido.

É possível conectar os conceitos dos autores anteriormente apresentados e inferir que até o início do Ensino Médio as avaliações precisariam enfatizar mais os processos mentais (*Gf/gp*) e, depois disso, deveriam medir preferencialmente o domínio do conhecimento (*Gc/Gkn/gk*), enquanto fator principal de diferença individual. E, deste modo, no presente estudo, a principal diferenciação entre um estudante bem sucedido (mais competente no domínio de sua área) e outro menos em tarefas de maior complexidade cognitiva, se refere principalmente ao domínio do conhecimento (Ackerman, 1996a, 1996b; Hambrick, 2005; Smith & Marsiske, 1994).

Também é importante destacar que, apesar de não ser o foco nesse estudo, sabe-se que outras variáveis importantes influem na aprendizagem e comprometimento dos estudantes. Nesse sentido, o interesse e as atitudes de uma pessoa sobre um domínio aumentam seu conhecimento no mesmo por meio do envolvimento em atividades relevantes. Independentemente da direção dessa influência, quanto mais uma pessoa sabe

sobre um domínio específico, mais fácil fica de adquirir conhecimentos adicionais sobre aquele domínio (Hambrick, 2005).

As perspectivas teóricas e empíricas apresentadas ao longo desse capítulo remetem a importância de se avaliar as habilidades e competências dos estudantes, a fim de verificar quais fatores (ou variáveis) são mais relevantes na diferenciação daqueles que poderiam ser chamados de aprendizes (novatos) e especialistas de um determinado domínio de conhecimento. Além disso, permitem reflexões sobre a importância da avaliação de características individuais envolvidas no processo de aquisição do conhecimento, bem como da necessidade de experiências adequadas de aprendizagem e do grau de dedicação aos estudos (também denominado na literatura como treino).

Em resumo, as contribuições de Ackerman são bastante coerentes com o objetivo do presente trabalho, pois destacam a importância de medidas diferentes para a avaliação dos estudantes ingressantes e concluintes na interpretação dos “produtos” da aprendizagem formal. Deste modo, entende-se que esse panorama teórico servirá de base para o melhor entendimento das variações de desempenho de ingressantes e concluintes avaliados pelo ENADE, e de outras variáveis associadas a esses momentos.

A pergunta sobre as melhores abordagens metodológicas para analisar as diferenças individuais na aquisição de conhecimentos e competências é certamente uma questão importante, e talvez a mais saliente, uma vez que todas as outras considerações teóricas diretamente ou indiretamente, dependem de um método adequado. No entanto, a controvérsia sobre a relação entre as habilidades e a aquisição de competências tem muitas facetas diferentes e ocorre em muitos níveis diferentes, por exemplo, quanto à natureza das tarefas de aprendizagem e suas características (por exemplo, a consistência da tarefa ou a complexidade da tarefa). Outro fator importante diz respeito à estrutura das capacidades

humanas e aquisição de habilidade visto que a maioria dos pesquisadores concorda que nem a aprendizagem nem as habilidades são realmente unitárias e, assim não devem ser mensuradas somente em termos correlacionais. Estes são pontos importantes que têm de ser tomados em consideração ao investigar os determinantes das diferenças individuais na aquisição de competências (Ackerman & Beier, 2005; Ackerman & Kanfer, 1993; Voelkle, Wittmann & Ackerman, 2006). Os autores também ressaltam a importância do controle de variáveis iniciais, medidas repetidas do mesmo sujeito e controle de outras variáveis a fim de determinar as diferenças individuais em termos de habilidades e competências.

Tais considerações são bastante relevantes quando se pensa na avaliação e indicadores provenientes dos resultados do ENADE. Embora o ENADE avalie o desempenho do estudante, em última instância, quando integrado aos outros indicadores do SINAES, pretende trazer informações sobre a qualidade dos cursos. Isso configura a principal questão sobre a validade do ENADE, isto é: Como é possível medir a qualidade de um curso por intermédio do desempenho de seus estudantes? Nesse exame são feitos os controles necessários das variáveis individuais, de contexto, entre outras, envolvidas nesse processo? Como é possível interpretar as pontuações no exame e os vários níveis de desempenho? Tais questionamentos são centrais no presente estudo visto que se a prova não for válida, não é possível inferir sobre as habilidades e competências dos estudantes. Diante dessas explanações, o próximo capítulo tem como objetivo discorrer sobre as interpretações geralmente divulgadas a partir de avaliações educacionais em larga como o ENADE, além das questões metodológicas e psicométricas relacionadas, que assim serão o foco no presente trabalho.

4. ESTUDO PSICOMÉTRICO DO ENADE

As informações dos estudos de validade e precisão podem trazer dados importantes para o aprimoramento de sistemas, na medida em que aprofundam o debate sobre a legitimidade e limitações das explicações dos resultados obtidos a partir dos instrumentos e apontam caminhos que devem ser seguidos para aumentar sua eficácia. Especificamente no contexto superior brasileiro, o ENADE pretende avaliar diferenças de influência entre instituições que oferecem o mesmo curso. E uma das interpretações decorrentes desse exame e que necessita ser validada, é que o resultado das diferenças entre estudantes concluintes e ingressantes simboliza os ganhos do estudante durante seu trajeto acadêmico (Primi, 2006b).

Apesar da importância dos estudos de validade para interpretação dos resultados e criação de medidas, e do impacto social decorrente das avaliações, existem poucos estudos feitos com os sistemas de avaliação desenvolvidos pelo MEC. Como o objetivo geral desse trabalho é estudar a validade do ENADE como elemento base para a avaliação das IES's e esse é um tema complexo, em seguida, serão apresentados três questionamentos psicométricos que serão tomados como centrais na investigação desse objetivo.

O construto medido pelas provas

Um dos pontos relacionados a validade do ENADE diz respeito às evidências sobre se, de fato, a prova avalia aquilo que se intencionou inicialmente no planejamento da mesma. Para tanto, uma das questões se refere à busca de evidências de validade baseadas na análise da estrutura interna das provas educacionais. Além disso, no presente estudo seria importante uma detalhada análise de conteúdo e dos componentes de dificuldade dos

itens da prova do ENADE a fim de proporcionar um melhor entendimento do construto mensurado, principalmente à luz da teoria CHC. Esses questionamentos são essenciais a fim de elucidar o que de fato se avaliou e quais as limitações das interpretações pretendidas, visto a implicação dos resultados decorrentes do exame nos indicadores de qualidade e nas políticas públicas (Primi, 2006b; Real, 2007; Schwartzman, 2010).

Como já mencionado, os componentes definidos nas diretrizes se constituem, para o ENADE, conteúdos a serem medidos. Assim, é necessário investigar o quanto a prova conseguiu operacionalizar a medida do construto, conforme se planejou inicialmente, e esses pontos se referem especialmente à abrangência do construto medido pelas provas (Pasquali, 2001; Verhine, Dantas & Soares, 2006),

Como também já argumentado, essa verificação da validade é muito relevante pois é praticamente impossível avaliar todos os construtos implicados nas diretrizes dos cursos com o formato e quantidade de itens propostos nas provas (Primi, 2006b, Primi & cols., 2009a, 2009b, Primi & cols., no prelo; Primi, Hutz & Silva, no prelo). Também é preciso analisar a existência e organização de todas as dimensões listadas, já que algumas delas podem estar relacionadas e representarem diferentes aspectos de uma mesma dimensão subjacente (Primi, 2006b).

Assim, os estudos de estrutura interna da prova para verificar quantas dimensões e de que forma se organizam na prova, comparando-se com as definições propostas, são de fundamental importância, dado o alcance limitado das provas. Os modelos mais usados nessas avaliações são baseados na TRI, mas somente se aplicam a situações em que haja uma unidimensionalidade essencial (Hambleton, Swaminathan & Rogers, 1991). A questão da definição das dimensões subjacentes à prova é relevante já que se as provas avaliarem mais dimensões não correlacionadas, a criação de uma única nota ignorando a estrutura

interna da prova poderá gerar uma série de problemas na interpretação desses escores (Primi & Nunes, no prelo).

Primi e Nunes (no prelo) e Primi, Hutz e Silva (no prelo) ressaltam um paradoxo interessante em relação a esse ponto uma vez que as comissões de especialistas produzem definições multidimensionais do que deve ser avaliado nas provas, enquanto que os especialistas em medida precisam construir provas unidimensionais para poder responder aos problemas de medida que enfrentam. De acordo com esses autores, embora o problema da unidimensionalidade e as questões correspondentes sejam de fundamental importância, poucos estudos estão voltados para o entendimento dessa questão (tais como, Laros, 2002; Condé & Laros, 2007; Rauter, Laros & Jesus, 2007).

Estudos como de Primi, Hutz e Silva (prelo), apontam uma tentativa de entendimento da existência de um fator geral em avaliações geralmente multidimensionais. Uma forma seria compreender a idéia da unidimensionalidade em termos do padrão de respostas esperados e a partir da organização dos itens em termos de complexidade das competências/habilidades/conteúdos avaliados. Dessa forma, a existência de um fator geral indicaria a sistematicidade nos perfis de acerto às questões em termos de domínio dos conhecimentos especializados da área. Tal entendimento, apesar de possibilitar uma interpretação diferenciada dos resultados, não soluciona a questão da unidimensionalidade, já que mesmo assim (como no ENADE) a prova mede bem menos do que proposto

A métrica das escalas

A mensuração e interpretação dos construtos são fundamentais para as pesquisas e práticas psicológicas e educacionais. Porém, o estabelecimento de significados dos intervalos escalares e as interpretações correspondentes às pontuações atribuídas é uma

questão corrente (Embretson, 2006). Para a interpretação de resultados de testes educacionais e psicológicos alguns métodos tem sido mais frequentemente utilizados, quais sejam, referência à norma, ao critério e ao conteúdo. A referência no desempenho do grupo normativo, apesar de ser o mais utilizado (inclusive no ENADE), visa à comparação do resultado de um indivíduo com os de um grupo-referência, não permite estabelecer o que a pessoa realmente é capaz de fazer. A referência no critério permite estabelecer parâmetros que diferenciem os sujeitos mais e menos habilidosos. E, por fim, a referência ao conteúdo é utilizada quando o conjunto de problemas presente no instrumento pode ser considerado uma amostra representativa do universo de problemas de um determinado conteúdo ou domínio (Primi, 2004; Cronbach, 1996).

A fim de comparar a avaliação referenciada na norma e referenciada no critério, Tamassia (2005) expõe alguns pontos de consideração, dos quais três se destacam. Em primeiro lugar, a finalidade do primeiro procedimento é classificar os estudantes, comparando o desempenho de cada estudante com os outros, enquanto que a finalidade do segundo procedimento é avaliar as aprendizagens individualmente, comparando o resultado do estudante com critérios ou padrões pré-estabelecidos. Em segundo lugar, a interpretação dos resultados na referência à norma se dá por meio de estatísticas relacionadas ao desempenho do grupo, enquanto que, na referência ao critério, a interpretação é dada por meio do desempenho do estudante em relação à obtenção ou não dos critérios definidos. Por fim, a distribuição dos resultados na referência normativa necessita, obviamente, seguir uma distribuição normal, e, então as questões da prova devem ser criadas de maneira a se obter o máximo de variabilidade de desempenho dos estudantes. Já na referência ao critério, a variabilidade dos resultados não é tão relevante, uma vez que as questões são criadas para se cobrir um conteúdo específico.

Semelhantemente, Woodcock (1999) destaca que o procedimento normativo apenas permite saber a posição que um indivíduo (ou qualquer outra unidade de análise que se considere no sistema, como a instituição de ensino) ocupa perante os outros, ou seja, se seu desempenho é superior ou inferior aos demais. Ele não possibilita afirmar com exatidão o que a pessoa é capaz de desempenhar, apenas o quanto ela realiza a mais ou a menos em relação à média. Assim, não implica, pelo menos diretamente, em afirmações precisas sobre a competência das pessoas. Além, disso, Bond (1996) ressalta que a referência na norma não permite dizer se o desempenho dos estudantes foi adequado ou não, apenas quais obtiveram desempenho melhor ou pior que os outros.

Por outro lado, o método referenciado no critério estabelece conceitos básicos de desempenho com os quais os estudantes são então comparados. Esse método, permite analisar um conteúdo mínimo que se espera que os estudantes tenham aprendido e sejam capazes de resolver numa prova. Outra diferença entre esses dois métodos diz respeito à interpretação do resultado final do estudante. Na referência à norma, os estudantes são distribuídos numa escala normal de desempenho, porém não necessariamente aqueles com maior desempenho obtiveram nota alta na prova. Eles podem ter acertado apenas metade da prova, ou menos, por exemplo, mas, ao serem comparados com todos os outros estudantes que acertaram ainda menos, seu desempenho fica parecendo elevado. Já na referência ao critério, tem-se um dado mais claro e fidedigno, uma vez que o parâmetro de comparação utilizado é o próprio conteúdo que se espera que os estudantes tenham aprendido durante o curso (Primi, Landeira-Fernandez & Ziviani, 2003).

De forma geral, Ferraz e cols. (1994) ressaltam que as perspectivas de interpretação criterial e normativa não são mutuamente exclusivas, pelo contrário, muitas vezes se inter-relacionam. Os autores destacam que são abordagens alternativas e, freqüentemente,

complementares, de maneira que as vantagens e desvantagens de cada uma delas dependem dos objetivos propostos nas avaliações, sendo ainda em função dessas finalidades que o avaliador se orienta em uma ou noutra direção.

Nesse sentido, o procedimento para interpretação com referência ao critério é considerado mais vantajoso nesse estudo pois pode possibilitar, por exemplo, o estabelecimento de pontos de corte, entre indivíduos mais e menos habilidosos, com o uso de critérios externos. Também Carroll (1993b) sistematizou um conjunto de procedimentos baseados na TRI, visando o melhor estabelecimento de significados para o desempenho, abordagem essa de maior interesse para o presente estudo. Esse autor chamou de escalonamento comportamental a busca de entendimento, em termos comportamentais, dos resultados dos testes com respeito ao que o examinado conhece ou pode realizar e assim, constitui-se como um aspecto da validade de construto.

O escalonamento comportamental englobaria um conjunto de métodos que visam dar um sentido mais amplo às interpretações feitas de um teste. Um dos aspectos do escalonamento comportamental se refere ao estabelecimento de padrões ou critérios para interpretação dos escores de um teste e esses são comumente designados na literatura como '*standard setting methods*'; e têm sido considerados fundamentais na apresentação dos resultados comparativos em muitos sistemas educacionais ao redor do mundo (Cizek, 1996; Impara & Plake, 2000; Shen, 2001; MacCann & Stanley, 2006).

O *Standard Setting* pode ser caracterizado pelo estabelecimento de um ou mais pontos de corte na escala de avaliação. Esse ponto de corte separa um nível de desempenho de outro, dividindo a distribuição do desempenho obtida pelos alunos em duas ou mais categorias distintas (Cizek & Bunch, 2007). Segundo esses autores, o estabelecimento desses pontos de corte deve ser definido de acordo com o tipo de categorias que se pretende

analisar. Caso seja necessário estabelecer somente uma diferenciação entre alunos proficientes ou não, é necessário somente um ponto de corte. No entanto, quando se pretende utilizar tal técnica com fins de medida de desempenho, é necessário acrescentar um número maior de pontos de corte para delimitar os diferentes níveis de desenvolvimento apresentado pelos testandos.

Segundo Ubriaco (2009) a determinação desses níveis de desempenho se baseia em três etapas. A primeira é a determinação de pontos de corte, que dividem a escala em categorias. A segunda é a estruturação dos Níveis de Desempenho, etapa que envolve a escolha do número de níveis, a seleção dos nomes dos níveis, a alocação dos itens nos níveis e a descrição pedagógica dos itens para caracterizar os níveis. E a terceira consiste na alocação das proficiências dos alunos, delimitando tais resultados em categorias/níveis. Essa autora levanta a importância de que as definições dos níveis sejam feitas a partir da análise dos itens pela TRI.

No entanto, segundo Kane (1994) é necessário distinguir a noção de pontos de corte e de nível de desempenho. O primeiro seria uma separação na escala, um ponto que divide a escala em diferentes partes. Já o nível de desempenho indica uma capacidade minimamente estabelecida, dentro de uma expectativa de desempenho na resolução de problemas que demandam o domínio de um determinado tipo de conhecimento e/ou habilidade. No presente estudo, os juízes devem estabelecer pontos de corte a partir dos níveis de desempenho no Componente Específico da prova do ENADE.

Outra possibilidade de escalonamento comportamental é a utilização de elementos que, além de avaliarem o conteúdo da matriz curricular presente nos itens, e sua dificuldade em termos de conhecimentos adquirido, avaliam também as habilidades cognitivas necessárias para o conhecimento, a compreensão e o pensar sobre um problema (Rodrigues,

1994). Esse tipo de análise de certa forma ultrapassa a análise dos juízes, pois exigem além do conhecimento na área que é avaliada, também conhecimento de processos cognitivos. Porém não pode prescindir do conhecimento dos juízes para sua realização, devendo ser realizada em conjunto. Tal procedimento aplicado às avaliações educacionais visa criar uma escala que possibilite estabelecer que aspectos foram aprendidos pelo aluno ao longo do curso. A nomeação desses níveis se dá de maneira a identificar a posição do aluno em relação a um desempenho almejado e planejado em relação à matriz curricular que ele deve aprender, e não somente em comparação com o desempenho de outros alunos.

Uma das formas para realizar o estabelecimento de pontos de corte se dá através da TRI em que é possível formar interpretações do significado das pontuações dos respondentes via referência aos itens que compõem o testes e é denominado, na literatura psicológica e educacional, como *Rasch Item Mapping Method*. Tal processo é feito a partir de um mapa de itens que estabelece a relação da dificuldade dos itens e o desempenho dos estudantes na prova (Embretson, 2006; Smith Jr., 2004).

Os métodos para estabelecimento de linhas ou pontos de corte mais utilizados parte do Método de Angoff e mais atualmente, do Mapa de Itens baseado no modelo de Rasch (como por exemplo, Cizek & Bunch, 2007; MacCann & Stanley, 2006; Shen, 2001), sendo o método de Angoff um dos mais tradicionalmente conhecidos e divulgados em exames de certificação de cursos, especialmente da área médica (Kaufman & cols., 2000; Searle, 2000). No método de Angoff é preciso que cada juiz julgue o nível de dificuldade de cada um dos itens que compõem um determinado teste e estime a probabilidade de um grupo hipotético de alunos em responder corretamente a cada uma das questões. Estas probabilidades são resumidas para obter a pontuação mínima atribuída por um único juiz. Logo, nesse método, o padrão de desempenho é estimado por meio da pontuação mínima

dada por cada especialistas da área e determinado pela somatória das estimativas para todos os itens de uma prova.

A interpretação criterial feita por meio do Mapa de itens parte da suposição de que a ordem de dificuldade dos itens calibrada pelo modelo reflete a complexidade dos conteúdos, em provas preponderantemente unidimensionais. Neste método, os especialistas têm a função de realizar uma análise qualitativa do desempenho dos estudantes a partir dos índices de dificuldade e habilidade estimados pelo modelo (Smith Jr., 2004). Assim, os especialistas precisam decidir qual é o nível mínimo de competência dentro da hierarquia dos conteúdos/dificuldades. Essa será a metodologia assumida no presente estudo a fim de proporcionar um melhor entendimento dos conteúdos, habilidades e competências avaliados no exame.

Shen (2001) realizou um estudo comparativo entre esses dois métodos a partir dos dados de um exame de admissão médico. Os resultados demonstram que, apesar dos padrões pelo método de Angoff terem sido mais severos, o mapa de itens refletiu melhor o critério desejado de competência mínima.

Oliveira, Franco e Soares (2007) realizaram a análise pedagógica dos itens do Projeto GERES por meio da TRI e nas curvas características dos itens consideraram como se dá a introdução, processamento e consolidação da habilidade. O principal ponto de destaque do artigo se refere à possibilidade de uma interpretação pedagógica e mais qualitativa dos parâmetros dos itens. Os autores ressaltam que tais resultados, provenientes da interpretação da escala de proficiência e do controle de outras variáveis intervenientes no processo de ensino e aprendizagem, contribuem para a prática educativa de maneiras diversas na medida em que oferece material para discussão do projeto pedagógico da escola, as formas de início, processamento e consolidação de habilidades.

Assim, um ponto importante para o aperfeiçoamento da validade das interpretações do ENADE será aplicar procedimentos de ajustamento de padrões baseados na TRI para se definir melhor os significados dos vários pontos da escala, além do que já se sabe com base nas informações normativas. Com isso, poder-se-á aclarar o que os diferentes níveis da escala significam em termos de acertos aos itens e, conseqüentemente, em termos dos construtos avaliados. Também se poderão interpretar melhor as diferenças encontradas entre cursos com diferentes conceitos e estudantes em diferentes momentos, a fim de compreender melhor os significados dos resultados, fazendo referência ao construto avaliado pela prova.

A medida da qualidade dos cursos

A avaliação da qualidade de cursos por meio do desempenho dos estudantes foi mais destacada quando se iniciou o Exame Nacional de Cursos (ENC) e tal procedimento está sujeito a limitações importantes referentes à validade (Landeira-Fernandez & Primi, 2002; Primi, Fernandez, & Ziviani, 2003; Raudenbush, 2004a; Soares, Castro & Ribeiro, 2000; Souza & Oliveira, 2003).

As principais críticas estão relacionadas ao uso do desempenho médio dos concluintes de um curso, comparado relativamente ao de outros cursos, como indicador direto de sua qualidade (Primi, 2006b, Primi & cols., 2009b, Primi & Nunes, no prelo; Primi, Hutz & Silva, no prelo). Entende-se que o perfil dos ingressantes de diferentes cursos é muito desigual visto que as notas dos concluintes "carregam" consigo a influência das diferenças prévias ao efeito que o curso produz no estudante, como, por exemplo, variáveis pessoais ou de contexto (Goldstein e cols., 1993; Goldstein, 2001; Soares e cols., 2001; Raudenbush, 2004). Também é importante destacar que como as diferenças de

desempenho prévias ao ingresso no curso não têm relação com o que o curso agrega e então a variância do desempenho no final, sem essa ponderação e correção, os resultados não podem ser usados como indicador da qualidade do curso (Primi, 2004; Primi, 2006b; Primi & Nunes, no prelo).

Nesse sentido, Brooke e Soares (2008) destacam que a ausência de mensuração do conhecimento prévio do aluno e de controle de outras variáveis não permite afirmações válidas e precisas sobre o efeito das escolas no processo de aprendizagem. Deste modo, um questionamento central frequentemente realizado nos estudos de efeito-escola diz respeito ao uso de medidas transversais e não longitudinais (Oliveira, Franco & Soares, 2007; Primi, 2006a; Primi & cols., 2009b; Rutter & cols., 2008a; Willms, 2008)

O ENADE pretende avaliar o efeito de diferentes cursos no desempenho dos estudantes e ordená-los em função do efeito que produzem. Primi (2006) porém destaca que há uma limitação importante presente em dados da avaliação da Educação Superior, como os do ENADE, que se refere a inexistência da alocação randômica dos estudantes aos diferentes cursos (Rubin, 1974). Esta alocação é chamada de randômica quando não há possibilidade alguma de se saber para qual grupo o próximo indivíduo será alocado. Assim, como o perfil dos ingressantes é bem heterogêneo entre os cursos, tanto em termos socioeconômicos quanto cognitivos (dentre outras variáveis), essas variáveis podem afetar o desempenho nas provas do ENADE, como apontado nas pesquisas anteriormente citadas.

Para uma interpretação mais precisa das diferenças no desempenho dos estudantes decorrentes do efeito do curso e da instituição freqüentada seria necessário a inserção de estudos experimentais (Primi, 2006b; Primi & Nunes, no prelo). Mas, embora isso seja impraticável no que se refere à avaliação em larga escala, como destacado por esses autores, é possível verificar em que medida dados como do ENADE permitem fazer

aproximações do efeito dos cursos sobre os estudantes, empregando controles estatísticos, tentando separar os efeitos dessas variáveis para se chegar a uma estimação do efeito da escola depreendido dessas variáveis confundidoras.

Rutter & cols. (2008a) corroboram tais questões ao realçar que outras metodologias devem ser adotadas para analisar as variáveis relacionadas ao funcionamento da escola e na diversificação das medidas de resultados. Para esses autores, a metodologia deve levar em consideração a escolha das medidas de desempenho que refletem nos reais aspectos do trabalho da escola, a necessidade da escola deter-se do conceito de organização social, incorporar a observação nas diferenças sociais entre as escolas à pesquisa e ainda identificar a aprendizagem e/ou outros atributos à atuação da escola.

A existência de fatores medidos em níveis distintos como estudante, escola e comunidade, afetando uma variável de resposta, nesse caso o desempenho do estudante, aponta a necessidade de uso de modelos de regressão multiníveis ou hierárquicos (Goldstein, 1995; Snijders & Bosker, 1999). O uso de tal metodologia é justificado não só por permitir a incorporação de fatores avaliados em níveis diversos, resolvendo o problema de unidade de análise, mas também pelo procedimento amostral utilizado que induz a uma estrutura hierárquica pela correlação entre as observações.

Além disso, como anteriormente apontado, alguns autores destacam que a base de estudo do desempenho escolar e das variáveis associadas para a avaliação do efeito-escola deve partir do controle de variáveis individuais (Ferrão, 2003; Ferrão & Fernandes, 2003). Assim, apesar de haver diferentes perspectivas a respeito de quais variáveis mais afetam o desempenho dos estudantes, é concordante o entendimento ao menos que o desenvolvimento cognitivo dos estudantes está condicionado às influências de

características extra-escolares (Brooke & Soares, 2008; Ferrão, 2003; Formiga, 2004; Soares, 2004) e o conhecimento que os estudantes já têm ao ingressarem.

Como exemplo de relevantes variáveis dos estudantes durante a trajetória acadêmica, levantadas por meio dos questionários socioeconômicos do ENADE, pode ser citado os resultados dos estudos desenvolvidos por Dias, Silva, Vendramini e Brito (2007) e Silva, Dias, Vendramini e Brito (2007). Essas pesquisadoras analisaram o desempenho dos estudantes dos vários cursos do ENADE 2004 e 2005 nas questões específicas de cada área e verificaram que os estudantes que indicaram um bom hábito de leitura e de estudo (Silva, Vendramini & Brito, 2007), além da participação em atividades de pesquisa e iniciação científica (Silva, Dias, Vendramini & Brito, 2007), apresentaram maiores notas no componente específico do exame.

Desta maneira, considera-se relevante investigar o quanto essas variáveis podem estar influenciando o desempenho dos estudantes, e como essas diferenças ocorrem entre ingressantes e concluintes. Nesse estudo, pretende-se decompor a variância do ENADE verificando que parcela é mais razoável estar associada à qualidade de curso e quais as potencialidades e limitações em se fazer medidas da qualidade dos cursos a partir das informações do desempenho dos estudantes (o quanto a influência de variáveis não escolares está sendo realmente ponderada na análise dos resultados).

5. OBJETIVOS

Objetivo Geral

O objetivo geral desse trabalho é discutir a validade e as implicações práticas e metodológicas em avaliar a qualidade dos cursos a partir de medidas do desempenho dos estudantes, como proposto no ENADE, por meio do modelo de Rasch e Multinível. Ou seja, até que ponto o desempenho do estudante reflete a qualidade de um curso e pode ser afirmado por meio dessas medidas.

Objetivos Específicos

Para tanto é preciso, em primeiro lugar, fazer um estudo de verificação de validade da estrutura interna da prova do ENADE para averiguar o número e organização das dimensões subjacentes aos itens e comparar com o que foi inicialmente planejado nas definições preparadas pelas comissões de área. Com isso, pretende-se descobrir quantas dimensões distintas de variabilidade a prova contém e os índices de precisão, dando subsídios psicométricos para a análise do construto medido pela prova.

Posteriormente, pretende-se compreender melhor as dimensões encontradas, empregando métodos da Teoria de Resposta ao Item, para o estabelecimento de critérios (Cizek, 1996; MacCann & Stanley, 2006), buscando associar significados mais claros aos pontos da escala, por intermédio da análise das variáveis definidoras dos itens e associadas às especificações da prova (conteúdos, habilidades e competências). Assim, esse procedimento pretende dar um sentido mais substancial aos diferentes pontos da escala, isto é, explicar melhor o que significam as variações nas dimensões encontradas. Nesse tipo de análise também serão usadas avaliações de juízes (especialistas) das áreas como parâmetros de probabilidade no modelo, de quais itens os estudantes mais competentes (*experts*)

deveriam ter mais chances de acertar, a partir da hierarquia dada pelo Método de Mapa de Itens. Esse estudo pode ser compreendido na classificação da literatura como evidência de validade pela análise do processo de resposta (AERA, APA & NCME, 1999).

Por fim, a fim de entender qual a variância relevante do desempenho do aluno que poderia carregar informações sobre a qualidade do curso, será realizado um estudo de validade por intermédio da verificação das relações com outras variáveis. Para tanto, pretende-se considerar o momento do percurso acadêmico e outras variáveis disponíveis do perfil prévio dos estudantes (provenientes do questionário socioeconômico) e das características dos cursos, como região, tipo de IES, entre outras para analisar o efeito dos cursos e instituições no desempenho dos estudantes.

A fim de atender aos três objetivos propostos para o presente trabalho e permitir que as informações sejam disponibilizadas de forma sistematizada, o método, resultados/discussão foram dispostos em três grandes grupos de estudos.

6. ESTUDO 1: Evidências de validade da estrutura interna da prova do ENADE

6.1 Método

Participantes

Os participantes provêm do banco de dados que contém os resultados dos estudantes do curso de Administração no ENADE de 2006, além das informações do questionário socioeconômico. O Banco de Dados do INEP estava disponível no Laboratório de Avaliação Psicológica e Educacional (LabAPE) da Universidade São Francisco, autorizado por meio do projeto aprovado pelo Edital 001/2006/ INEP/ CAPES - Observatório da Educação. Esse curso foi selecionado por conveniência a fim de ilustrar a proposta de análise realizada e dar uma ênfase maior a avaliação qualitativa dos resultados alcançados. Além disso, trata-se de um curso em que houve uma maior disposição dos juízes e cujo conteúdo e área de atuação profissional faz interface com diferentes outros cursos não somente da área de humanas. Foram considerados para a análise somente os estudantes selecionados e que haviam respondido as duas partes da prova, sendo 55% deles ingressantes (Tabela 1).

Tabela 1 - Distribuição de estudantes selecionados no ENADE 2006 de Administração

Estudante	Número de estudantes	%
Ingressante	71.838	55,1
Concluinte	58.428	44,9
Total	130.266	100,0

Do total de participantes, 52% são do sexo feminino, com idade média de 26 anos e desvio padrão de 7. Em relação à região de funcionamento do curso, aproximadamente a metade dos estudantes estudou na região sudeste (49%), seguido da Sul (22%) e Nordeste (14%). Das instituições de ensino participantes nesse ano do ENADE, a grande maioria provém da rede particular de ensino (90,2%). A Tabela 2 mostra a distribuição de estudantes por organização acadêmica, em que aproximadamente 52% vêm de Faculdades, Escolas e Institutos Superiores.

Tabela 2 - Distribuição de estudantes no ENADE 2006 de Administração por organização acadêmica

Organização acadêmica	Número de estudantes	%
Universidade	38.665	29,7
Centro Universitário	16.367	12,6
Faculdades Integradas	7.825	6,0
Faculdades, Escolas e Institutos Superiores	67.285	51,7
Centros de Educação Tecnológica	124	0,1
Total	130.266	100,0

Materiais

Banco de dados do Exame Nacional do Desempenho do Estudante (ENADE)

A prova do ENADE apresenta 40 questões que avaliam conhecimentos, competências profissionais e habilidades acadêmicas de estudantes do Ensino Superior. Dessas questões, 10 são relativas ao componente de Formação Geral (sendo 8 objetivas e duas dissertativas), e 30 itens específicos da área do curso em avaliação (componente específico avaliado por 24 questões objetivas e seis dissertativas).

Questionário Socioeconômico

Além do ENADE, os estudantes responderam a um questionário socioeconômico com 114 questões, indagando sobre perfil socioeconômico, dedicação às atividades acadêmicas, hábitos de leitura e conhecimento de línguas, história acadêmica, uso do computador e acesso à internet, e percepção de elementos da instituição e da formação por ela oferecida. Por intermédio desse questionário, é possível, extrair dimensões importantes a respeito do perfil dos estudantes, além de obter dados contextuais característicos dos diferentes cursos. Esse questionário é preenchido pelos alunos em casa e entregue no dia da prova, e ao examinar a quantidade de participantes nessa avaliação verificou-se que 80% dos estudantes responderam a esse instrumento.

Além desses instrumentos, os bancos de dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) mantêm informações gerais sobre os cursos, tais como, número de estudantes, dependência administrativa, dentre outras, que serviram para caracterizar os cursos e comparar os resultados.

Procedimento

Após a autorização para a utilização dos bancos de dados e aprovação do projeto pelo Comitê de Ética (Anexo 1), os dados foram organizados de forma a conter as informações somente dos estudantes que haviam sido selecionados para o exame e com padrões de resposta lineares (sem muitas respostas em branco, e com pelo menos a parte de Componente Específico respondido).

6.2 Resultados e Discussão

Nessa fase, foi usado o programa *Winsteps* para calibração dos itens (Linacre & Wright, 1991). As questões discursivas recebem originalmente pontuação não dicotômica que varia de 0 a 100 pontos. Para esse estudo, essas notas foram transformadas para uma pontuação de 1 a 9 pontos a fim de facilitar a interpretação dos resultados e a leitura do programa estatístico.

O programa permite a estimação dos parâmetros pela TRI de acordo com o modelo de Rasch, em que se busca definir a métrica da escala geral subjacente aos itens e o nível de habilidade em uma escala de valores *theta* (θ), que graduam as habilidades avaliadas pelo teste (Wright & Stone, 1979). Esse modelo é usado para agregar dados de respostas aos itens calibrando-se medidas intervalares para os itens (nível de dificuldade) e para os estudantes (nível de competência). Como no ENADE misturam-se questões de múltipla escolha com questões dissertativas foi necessário analisar os dados com o modelo de créditos parciais, modelo mais recente na família de modelos de Rasch, que permite a calibração simultânea de itens com diferentes estruturas dicotômicas e politômicas. A calibração conjunta permite que se obtenham as medidas dos sujeitos na mesma escala.

Foi aplicada a TRI para estimação dos parâmetros dos itens e das pessoas nas duas partes da prova, de acordo com o modelo de Rasch (por meio do programa WINSTEPS, Linacre, & Wright, 1991). Na Tabela 3 estão apresentados os resultados gerais das medidas dos estudantes e dos itens na parte de Formação Geral. Dessas informações, as mais importantes são as seguintes: (a) as medidas dos estudantes (*habilidade*) variaram de um mínimo de -2,80 a um máximo de 3,04 com média -0,29 e desvio padrão 0,72; (b) a precisão geral dessas medidas foi de 0,41 nível baixo porém aceitável devido a quantidade de itens nessa parte da prova; (c) o índice de separação para a medida dos estudantes foi de

0,84 indicando precisão estatística para separar de um grupo de estudantes, valor considerado pouco discriminativo. O índice de separação é calculado pela razão da variância verdadeira (corrigida pela estimativa de variância de erro) com o erro de medida, indicando a dispersão das medidas dos estudantes em unidades de erro padrão, isto é, quantas vezes a dispersão verdadeira entre as medidas supera o erro padrão da medida; (d) as medidas de desajuste dos itens (INFIT e OUTFIT) estiveram em níveis aceitáveis atingindo no máximo 1,25. Em resumo, considerando a pequena quantidade de itens nessa parte da prova, esses dados indicam um bom ajuste dos valores calibrados pelo modelo com os dados observados. Optou-se pela não seleção/exclusão de alguns itens como tentativa de melhorar a medida pois a intenção da presente análise era de averiguar a qualidade dos itens da prova como ela então foi aplicada.

Tabela 3 - Estatísticas gerais dos estudantes e itens no ENADE 2006 de Administração segundo o modelo de Rasch na parte de Formação Geral

	Estatísticas	Escore bruto	Itens válidos	Habilidade	Erro modelo	Infit	Outfit
Estudantes	Média	12,5	9,7	-0,29	0,48	0,87	1,04
	DP	4,9	0,7	0,72	0,15	0,81	0,74
	Máximo	25,0	10,0	3,04	1,70	9,90	9,90
	Mínimo	1,0	1,0	-2,80	0,40	0,00	0,00
	Separação real		0,84		Precisão real		0,41
	Separação modelo		1,02		Precisão modelo		0,51
	Erro padrão da média		0,00				
	Estatísticas	Escore bruto	Casos válidos	Dificuldade	Erro modelo	Infit	Outfit
Itens	Média	139562,1	125261,0	0,00	0,01	0,99	1,04
	DP	176881,4	7347,7	1,17	0,00	0,11	0,14
	Máximo	589980,0	129103,0	2,06	0,01	1,10	1,25
	Mínimo	13351,0	108757,0	-1,84	0,00	0,69	0,73
	Separação real		180,7		Precisão real		1,00
	Separação modelo		184,0		Precisão modelo		1,00
	Erro padrão da média		0,39				

Na Tabela 4 estão apresentadas as estatísticas dos itens de Formação Geral. Dessas, as estatísticas mais importantes são o valor estimado da dificuldade b de cada item (*Measure*), os valores de ajuste por item (INFIT e OUTFIT) e as correlações item-theta. Verifica-se que a prova foi relativamente mais difícil já que a diferença habilidade/dificuldade foi -0,29. Esses dados mostram que todos os itens apresentam ajustes aceitáveis a despeito de alguns possuírem baixa correlação item-theta (Itens FG5 e FG7). Esses itens também foram os que apresentaram os índices OUTFIT mais elevados.

Tabela 4 - Estatísticas dos itens em Formação Geral no ENADE 2006 de Administração após calibração pelo modelo de Rasch

Item	Score Bruto	N	Dificuldade	Erro do Modelo	Infit	Outfit	Correlação item-theta	ID
FG_07	13.351	128.979	2,06	0,01	1,05	1,25	0,00	0,10
FG_05	28.770	128.955	1,09	0,01	1,10	1,22	-0,02	0,22
FG_01	30.146	128.896	1,02	0,01	1,04	1,08	0,08	0,23
FG_04	35.722	128.838	0,77	0,01	1,04	1,07	0,08	0,28
FG_D9	355.313	112.576	0,11	0,00	0,69	0,73	0,20	4,15
FG_06	68.596	128.722	-0,44	0,01	0,99	0,98	0,17	0,53
FG_02	74.564	129.103	-0,64	0,01	1,06	1,07	0,07	0,58
FG_D10	589.980	108.757	-0,67	0,00	0,99	1,07	0,18	6,42
FG_03	95.700	128.831	-1,46	0,01	1,00	0,99	0,15	0,74
FG_08	103.479	128.953	-1,84	0,01	0,96	0,92	0,21	0,80
Média	139.562	125.261	0,00	0,01	0,99	1,04		
DP	176.881	7347,7	1,17	0,00	0,11	0,14		

Conforme dito anteriormente, uma vantagem do modelo de Rasch é que as estimativas de dificuldade dos itens e competências dos estudantes têm uma métrica comum. Portanto, com o valor da medida de competência do estudante é possível prever quais itens ele terá mais chances de acertar e errar, dependendo de sua posição na escala.

Tabela 5 - Probabilidade de acerto esperada

Habilidade - Dificuldade	Chance de Acerto (%)
-3,0	1
-2,5	1
-2,0	3
-1,5	7
-1,0	15
-0,5	30
0,0	50
0,5	70
1,0	85
1,5	93
2,0	97
2,5	99
3,0	99

O mapa dos itens da parte de Formação Geral da prova de Administração do ENADE 2006 é apresentado na Figura 3. Nessa Figura a escala das medidas das competências dos sujeitos é desenhada na linha horizontal (na parte inferior da figura) variando de -3 a $+4$. Os valores T, S, M indicam o ponto em que se encontra a média (M), um desvio padrão (S) e dois desvios padrão (T) das medidas de competência dos 129.802 estudantes. Os números acima dessas letras indicam o número de estudantes em cada nível da escala, por exemplo, a partir deles pode-se dizer que 11.214 estudantes tiveram thetas iguais a média ($-0,29$).

De cima para baixo, na coluna abaixo do theta de valor 4 estão as alternativas de resposta consideradas certas no caso dos itens objetivos e o padrão de resposta no caso das questões dissertativas, em que um estudante altamente competente deveria acertar com nota 9. O ponto “:” nas questões objetivas marca o ponto em que o acerto ao item passa a ser mais provável que o erro (índice de dificuldade). E nas questões dissertativas esses pontos variam de 1 a 9 e indicam os padrões de resposta de acordo com o aumento de habilidade, ou seja, indicam o nível de habilidade em que se espera a mesma probabilidade do sujeito

responder o item abaixo ou acima. Os itens estão ordenados de baixo para cima em ordem crescente de dificuldade. Pode-se observar, por exemplo, que no item mais fácil, o acerto passa a ser mais provável partir do nível -1,8 (dificuldade do item). Se considerado o nível -0,9 e colocada uma linha perpendicular nesse ponto até a linha superior pode-se observar o cruzamento dessa linha com os escores esperados (buscando em quais itens a linha passa ao lado direito do ponto “:”) e prever que os estudantes nesse nível de competência acertariam os quatro itens mais fáceis e teriam escore 6 no item 11, dissertativo.

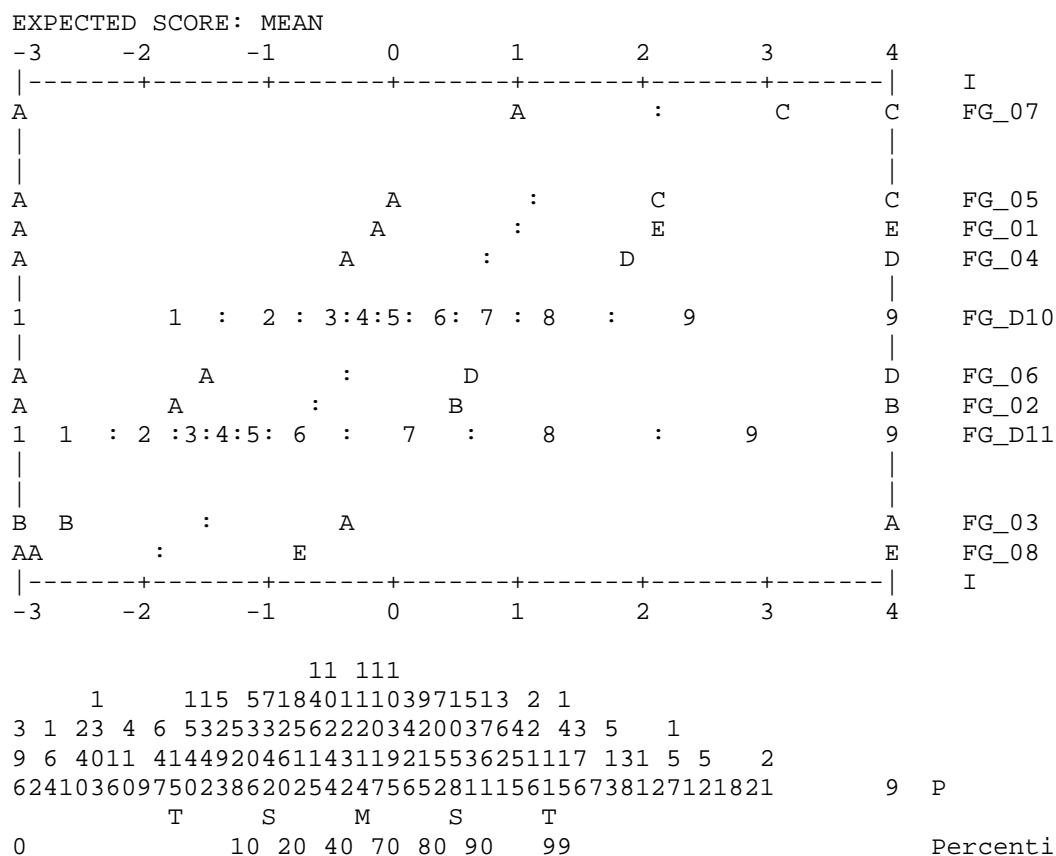


Figura 3 - Mapa dos itens da prova Formação Geral do ENADE 2006 de Administração

Na Tabela 6 estão apresentados os resultados gerais das medidas dos estudantes e dos itens no Componente Específico. As medidas dos estudantes (*habilidade*) variaram de

um mínimo de -3,71 a um máximo de 2,26 com média -0,61 e desvio padrão 0,53; a precisão geral dessas medidas foi de 0,56 nível aceitável devido ao propósito da prova; o índice de separação para a medida dos estudantes foi de 1,13 indicando que pode ter de um a dois estratos de escores para separar com precisão os estudantes, mas valor considerado pouco discriminativo. As medidas de desajuste dos itens (INFIT e OUTFIT) estiveram em níveis aceitáveis atingindo no máximo 1,12. Em resumo, considerando a quantidade de itens nessa parte da prova, esses dados indicam um bom ajuste dos valores calibrados pelo modelo com os dados observados.

Tabela 6 - Estatísticas gerais dos estudantes e itens no ENADE 2006 de Administração segundo o modelo de Rasch na parte de Componente Específico

	Estatísticas	Escore bruto	Itens válidos	Habilidade	Erro modelo	Infit	Outfit
Estudantes	Média	22,0	27,9	-0,61	0,31	1,02	1,02
	DP	10,9	2,5	0,53	0,12	0,50	0,33
	Máximo	72,0	30,0	2,26	1,59	9,90	8,19
	Mínimo	1,0	1,0	-3,71	0,20	0,00	0,00
	Separação real		1,13		Precisão real		0,56
	Separação modelo		1,27		Precisão do modelo		0,62
	Erro padrão da média		0,00				
	Estatísticas	Escore bruto	Casos válidos	Dificuldade	Erro modelo	Infit	Outfit
Itens	Média	77416,6	120717,6	0,00	0,01	1,00	1,01
	DP	85257,1	16738,6	0,88	0,00	0,04	0,07
	Máximo	398422,0	129519,0	1,67	0,01	1,12	1,12
	Mínimo	13024,0	80587,0	-2,05	0,00	0,93	0,83
	Separação real		144,7		Precisão real		1,00
	Separação modelo		145,7		Precisão do modelo		1,00
	Erro padrão da média		0,16				

Por esses dados, identifica-se que a prova foi relativamente mais difícil já que a diferença habilidade/dificuldade foi -0,61. Na Tabela 7 estão apresentadas as estatísticas dos itens do Componente Específico. Esses dados mostram que todos os itens apresentam ajustes aceitáveis a despeito de alguns possuírem baixa correlação item-total, que foram

também os itens que apresentaram os índices OUTFIT mais elevados (acima de 1,05). Essas informações das Tabelas 6 e 7 mostram o padrão ajustado dos estudantes e itens, e indica igualmente a existência de uma primeira dimensão (posteriormente serão apresentados os resultados da análise fatorial de resíduos).

Tabela 7 - Estatísticas no Componente Específico no ENADE 2006 de Administração após calibração pelo modelo de Rasch

Item	Escore Bruto	N	Dificuldade	Erro do Modelo	Infit	Outfit	Correl. item-theta	ID
CE_31	13024	128.514	1,67	0,01	1,02	1,12	0,02	0,10
CE_34	15931	127.715	1,43	0,01	1,01	1,10	0,04	0,12
CE_17	21647	129.393	1,08	0,01	0,99	1,01	0,12	0,17
CE_D37	17292	94.865	0,96	0,01	0,99	0,83	0,23	1,18
CE_23	26442	129.019	0,82	0,01	1,04	1,10	0,01	0,20
CE_21	28695	129.362	0,72	0,01	1,04	1,10	0,02	0,22
CE_33	29647	128.221	0,66	0,01	1,03	1,09	0,04	0,23
CE_D35	80924	90.582	0,66	0,00	1,02	1,03	0,30	1,89
CE_32	30711	128.456	0,62	0,01	1,04	1,11	0,01	0,24
CE_11	31031	129.314	0,61	0,01	1,00	1,01	0,11	0,24
CE_26	36559	128.540	0,37	0,01	0,99	1,00	0,14	0,28
CE_27	39408	129.226	0,27	0,01	1,06	1,11	0,02	0,30
CE_22	42528	129.395	0,15	0,01	1,00	1,01	0,12	0,33
CE_D36	131514	85.039	0,14	0,00	1,03	1,03	0,33	2,55
CE_29	44240	129.125	0,09	0,01	0,97	0,97	0,19	0,34
CE_25	44862	128.850	0,06	0,01	0,99	0,99	0,15	0,35
CE_24	45622	128.457	0,03	0,01	0,93	0,92	0,26	0,36
CE_D39	181046	80.587	-0,06	0,00	1,02	1,01	0,34	3,25
CE_30	50034	128.795	-0,12	0,01	0,95	0,94	0,24	0,39
CE_20	52831	129.362	-0,21	0,01	0,98	0,99	0,17	0,41
CE_12	53976	129.316	-0,25	0,01	1,00	1,01	0,12	0,42
CE_D40	330688	92.207	-0,34	0,00	1,12	1,12	0,24	4,59
CE_28	60924	128.861	-0,49	0,01	0,98	0,97	0,20	0,47
CE_19	62411	129.396	-0,53	0,01	0,99	0,98	0,16	0,48
CE_D38	398422	82.233	-0,74	0,00	1,01	1,03	0,34	5,84
CE_18	84711	129.303	-1,29	0,01	0,97	0,97	0,18	0,66
CE_16	87201	129.519	-1,37	0,01	1,04	1,08	0,04	0,67
CE_13	87952	129.229	-1,41	0,01	0,94	0,92	0,23	0,68
CE_15	89311	129.324	-1,46	0,01	0,96	0,96	0,17	0,69
CE_14	102914	129.322	-2,05	0,01	0,93	0,87	0,23	0,80
Média	77416.6	120.717	0,00	0,01	1,00	1,01		
DP	85257.1	16.738	0,88	0,00	0,04	0,07		

Ao analisar os itens do CE após a calibração, considerando os dados tanto dos estudantes ingressantes quanto concluintes, especialmente aqueles com correlação acima de 0,20 que então estariam mais associados com o Fator Geral, verifica-se a concentração de itens dissertativos. Além disso, dentre os itens objetivos, encontram-se itens também de maior complexidade e/ ou cuja resolução necessita da leitura e interpretação de gráficos e tabelas. Finalmente, o mapa dos itens do Componente Específico da prova de Administração do ENADE 2006 é apresentado na Figura 4.

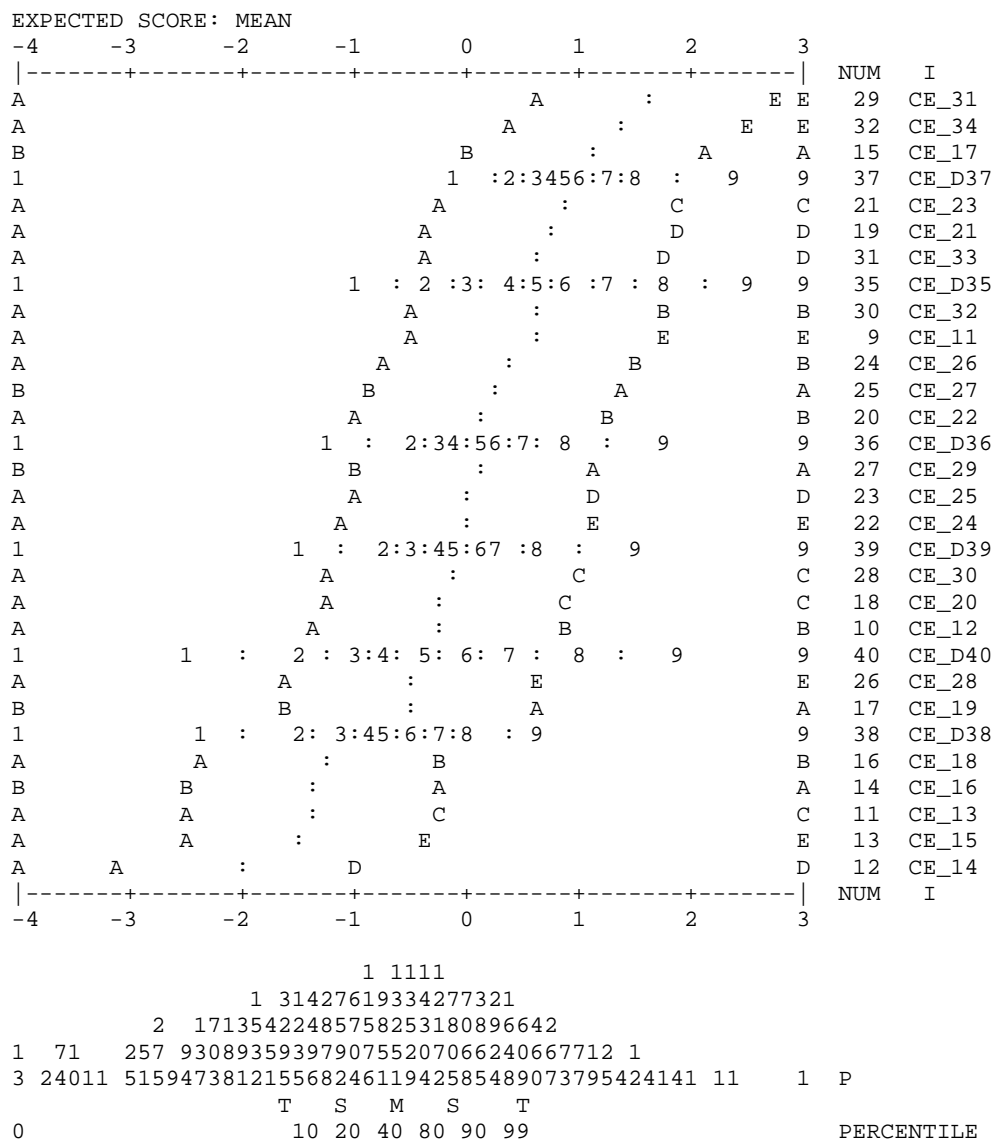


Figura 4 - Mapa dos itens do Componente Específico do ENADE 2006 de Administração

A escala das medidas das competências dos sujeitos variou de -4 a $+3$ quando mensurada a competência dos 129.709 estudantes. Verifica-se que 13.251 estudantes tiveram thetas iguais a média ($-0,61$). Pode-se observar, por exemplo, que no item mais fácil, o acerto passa a ser mais provável partir do nível -2 . Se considerado o nível -1 é possível prever que os estudantes nesse nível de competência acertariam os seis itens mais fáceis (CE14, CE15, CE13, CE16 e CE18) e teriam score 3 no item 38, dissertativo.

Após considerar a estatística dos itens, a disposição dos mesmos no mapa e a análise qualitativa conjunta com os juízes, verifica-se que em geral, a dimensão principal da prova envolve conhecimentos, habilidades e competências em Teorias da Administração, Administração Financeira e de Gestão estratégica. Ou seja, conhecimentos básicos e complexos voltados para a prática profissional.

A fim de buscar evidências de validade da estrutura interna dos itens e relacionar com as especificações iniciais, realizou-se a análise dos componentes dos resíduos para verificar a estrutura interna formada a partir de suas correlações. Assim, pretendeu-se examinar a dimensionalidade da prova investigando se os itens avaliam uma ou mais dimensões e comparar com as especificações definidas na matriz (diretrizes) de avaliação do ENADE.

Primeiramente procedeu-se a análise da dimensionalidade da prova referente ao Componente Específico (CE) por meio da análise de componentes dos resíduos. Essa análise estuda correlações de resíduos que são entendidos como respostas inesperadas e indica a co-ocorrência desses em um subconjunto de itens, indicando a possível existência de outras dimensões além da primeira. Para tanto, o programa efetua primeiramente a calibração da dificuldade dos itens e da habilidade das pessoas; a partir disso são calculadas as respostas esperadas; é feito o cálculo dos resíduos (resposta observada-esperada) e a

preparação da matriz residual (em que se remove a influência da primeira dimensão); calculam-se as correlações entre esses resíduos para então permitir a análise de componentes principais dessa matriz de correlação entre resíduos

Os componentes dos resíduos para os testes de inteligência, habilidade e conhecimentos mostram os contrastes, ou seja, a formação de agrupamentos distintos de itens, geralmente indicativo de dimensões secundárias e/ou mau funcionamento de alguns itens. Na análise fatorial convencional, a interpretação pode ser baseada apenas em cargas positivas. No julgamento dos resíduos, a interpretação deve ser baseada no contraste entre cargas positivas e negativas. Como o programa mostra os resultados dos componentes de resíduos, o resultado do primeiro componente (dimensão) já foi removido e passa-se a analisar as dimensões secundárias, ou também chamados de componentes dos contrastes. Ao interpretar o significado de um componente ou um fator, a abordagem convencional sugere olhar para as maiores cargas positivas, a fim de inferir o significado substantivo do componente. Na análise de resíduos, é necessário identificar os agrupamento de resposta nos contrastes e interpretar o significado dos componentes. O primeiro contraste pode tanto representar outra(s) dimensão (ões) dependendo dos valores encontrados (se *eigenvalues* maiores do que 2,0), ou mesmo uma mudança de intensidade do construto principal (Schölkopf, Smola & Müller, 1999; Smith & Smith, 2004; Wright & Stone, 1979).

Ao analisar os resultados dos componentes principais dos resíduos do componente específico do ENADE 2006 de Administração, de acordo com o programa *Winsteps* (Tabela 8) verifica-se que a prova pode ser considerada preponderantemente unidimensional de forma que 54,4% da variância dos dados podem ser explicadas pela dimensão principal. E, apesar da indicação de cinco componentes principais dos resíduos (contrastos), a maior dimensão secundária, ou seja, o primeiro contraste nos resíduos

explica somente 2,3% da variância. Os *eigenvalues* dos cinco primeiros fatores foram (1,5; 1,3; 1,2; 1,2 e 1,1) sugerindo somente a intensificação do construto principal avaliado pela prova e não a presença de outras dimensões ou construtos de forma substancial.

Os itens com cargas superiores a 0,20 do primeiro contraste positivo e negativo podem ser visualizados na Tabela 8. Optou-se no presente estudo por esclarecer mais qualitativamente a respeito somente do primeiro contraste, por considerá-lo como um dos principais elementos para discussão dos resíduos encontrados.

Tabela 8 - Distribuição fatorial e de dificuldade dos principais itens nos contrastes

Contraste 1A (positivo)			Contraste 1B (negativo)		
Item	Carga	Dificuldade	Item	Carga	Dificuldade
CE_13	0,45	-1,41	CE_D40	-0,41	-0,34
CE_24	0,42	0,03	CE_27	-0,33	0,27
CE_30	0,41	-0,12	CE_32	-0,27	0,62
CE_14	0,37	-2,05	CE_21	-0,22	0,72
CE_18	0,23	-1,29	CE_33	-0,20	0,66
CE_15	0,20	-1,46	CE_16	-0,19	-1,37
CE_25	0,19	0,06			

No primeiro contraste, positivo encontram-se itens classificados como fáceis e médios. As habilidades que se objetivava avaliar, de acordo com as diretrizes da prova, em geral dizem respeito a: Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1), Articulação do conhecimento sistematizado com a ação profissional (2) e Raciocínio lógico na identificação de problemas organizacionais (3). No primeiro contraste, negativo encontram-se mais itens classificados como médios. Em relação às habilidades, foram aferidas aquelas relacionadas ao primeiro fator positivo, além de: Hierarquização de atividades e programas para tomada de decisões

(9), Percepção de modelos de estratégia e seu impacto na gestão (10), Entendimento dos canais de distribuição e cadeia de suprimento (12).

Algumas das questões principalmente do contraste positivo foram também apontadas pelos juízes (como será visto no segundo estudo) como envolvendo conteúdos básicos porém cujas alternativas se apresentam confusas ou por mensurarem mais interpretação de texto do que domínio específico de um conhecimento. Já no contraste negativo, encontram-se questões em que é exigida maior capacidade de abstração e de aplicação prática do conhecimento, interpretação de gráficos, figuras e tabelas; ou seja, envolvem maior complexidade do conhecimento.

É interessante analisar o conteúdo dos itens com carga fatorial acima de 0,30 (CE13, CE14, CE24, CE27, CE30, CE32 e CE40) pois esses podem dar um indicativo de como interpretar o sentido da segunda e terceira variável latente em termos de competências/habilidade/conteúdos avaliados. Esses itens, em termos de conteúdos previstos nas diretrizes, tratam de Planejamento; Teorias da Administração; Administração de Recursos Humanos e Psicologia; Administração Financeira e Orçamentária; Contabilidade; Administração de Serviços; Administração de Produção; Processos; Logística; Estatística e Matemática. Também a partir das Tabelas 9 e 10 é possível verificar a distribuição de conteúdos predominantes e habilidades que se gostaria de aferir por meio desses itens (de acordo com diretrizes da prova) dos itens apontados na segunda e terceira variável latente.

O que se pode verificar, complementando o que já foi apontado anteriormente, é que na dimensão secundária positiva (segunda variável latente), qualitativamente avaliando, esses itens, em geral, envolvem a capacidade de interpretar textos e informações dispostas em forma de figuras, tabelas e gráficos. Além disso, algumas questões foram apontadas como confusas pelos juízes, quando analisados os distratores.

Tabela 9- *Conteúdos predominantes, habilidades e competências dos itens do contraste positivo (segunda variável latente)*

Item	Conteúdos predominantes	Habilidades aferidas
13	Teorias da Administração, Planejamento	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Compreensão do todo administrativo e de suas relações com o ambiente externo (4)
14	Teorias da Administração, Administração de Recursos Humanos e Psicologia.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)
24	Administração Financeira e Orçamentária, Contabilidade e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)
30	Administração de Produção, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Resolução de problemas organizacionais com flexibilidade e adaptabilidade (8)

A dimensão secundária negativa (terceira variável latente) envolve a resolução de itens um pouco mais complexos e, pode-se inferir também que esses itens requerem a competência de mobilizar os conhecimentos e aplicá-los em uma situação problema ou contexto mais voltado à prática profissional, especialmente no item dissertativo.

Tabela 10- *Conteúdos predominantes, habilidades e competências dos itens do contraste negativo (terceira variável latente)*

Item	Conteúdos predominantes	Habilidades aferidas
27	Administração de Serviços	<ul style="list-style-type: none"> ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na solução de problemas organizacionais (3) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos da organização (7) ▪ Tomada de decisão em serviços (9) ▪ Percepção de modelos de gestão de serviços (10)
32	Administração de Produção, Processos, Logística, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Hierarquização de atividades e programas para tomada de decisões (9) ▪ Percepção de modelos inovadores de gestão (10) ▪ Compreensão dos processos produtivos e controle estatístico dos processos (11) ▪ Entendimento dos canais de distribuição e cadeia de suprimento (12)
40	Teorias da Administração, Administração de Recursos Humanos, Psicologia e Sociologia	<ul style="list-style-type: none"> ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Operação com formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais. ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)

Apesar das diretrizes apontarem para 8 habilidades e competências gerais, ao verificar as especificações de cada um dos itens no Componente Específico da prova observa-se que na verdade são propostas as avaliações de 14 habilidades específicas, que foram assim agrupadas apesar de algumas diferenças nas denominações (ora mais gerais, ora mais detalhistas), como é mostrado no Anexo 2. Além disso, verifica-se que essas são referenciadas como habilidades e competências, e em alguns momentos somente como habilidades, demonstrando assim a dificuldade de separação entre o que é considerado habilidade e o que é competência.

Em geral, a predominância unidimensional mostrada na análise do programa pode ser contraditória ao pensar na quantidade de conteúdos, habilidades e competências referenciados pelo INEP para a compreensão dos itens, o que daria a noção de uma prova multidimensional. Além disso, outra característica a se criticar e já apontada por outros autores (Primi, 2006b; Primi & cols., 2009a, 2009b; Primi & cols., no prelo; Primi, Hutz & Silva, no prelo; Verhine, Dantas e Soares, 2006) refere-se ao questionamento a respeito da quantidade de itens versus habilidades que se propõe mensurar, de forma que se pode verificar que em cada um dos itens é proposto aferir muitas habilidades e nem todas as habilidades são devida e linearmente distribuídas ao longo da prova (ver quadro do Anexo 2). No entanto, estudos como de Primi, Hutz e Silva (no prelo), sugerem ser mais fácil compreender a idéia da unidimensionalidade a partir da organização dos itens, ordenando-os em uma linha em função de sua dificuldade (essa hierarquia será mais detalhada adiante na análise do mapa de itens da prova). Os itens têm diferentes níveis de dificuldade e isso ocorre porque há diferenças na complexidade das competências/habilidades/conteúdos avaliados. Assim, a existência de um fator geral pode ser entendida pelo padrão de resposta dos estudantes, em que aqueles que acertam os itens complexos tendem a acertar os itens

anteriores, considerados como mais fáceis na hierarquia. Portanto, segundo esses autores, uma possibilidade de entendimento do fator geral apesar das diretrizes multidimensionais, seria a indicação dessa “sistematicidade” nos perfis de acerto às questões, indicando também uma organização progressiva da aquisição das competências/habilidades/conteúdos principais representadas pelas questões. E, a presença de um fator geral subjacente, nessa perspectiva, indicaria que os padrões esperados são mais frequentes do que os inesperados.

Porém, tal interpretação ainda não soluciona uma das questões metodológicas discutidas a respeito do exame (Primi, Hutz & Silva, no prelo; Primi e cols., 2009b; Primi, 2006b) que se refere a validade em avaliar as competências profissionais e habilidades acadêmicas conforme planejado nas diretrizes com a quantidade de itens propostos e quais os parâmetros referenciais para sua interpretação. Tal questionamento está relacionado à limitação do número dos itens que impede que se consiga avaliar com nível alto de confiabilidade diferentes elementos do construto (aqui entendido como as graduações ou vários domínios do conhecimento).

Portanto, essa dimensão mais global se refere ao domínio de alguns conhecimentos específicos, mais complexos e imprescindíveis para a prática profissional, porém mais restrita ao que está especificado nas diretrizes. Também, como já apontado, tal resultado sobre a predominância de um fator geral referente ao domínio de alguns conhecimentos específicos, especialmente aqueles sobre Teorias da Administração (ênfatisados principalmente até a metade do curso) e Administração Estratégica e Mercadológica (dois últimos anos de curso), apesar das diretrizes multidimensionais da prova, permite a inferência de que apesar da variação em termos de conteúdos e habilidades nas diretrizes, no geral, avaliações educacionais como o ENADE tem como base subjacente principalmente o grau de domínio dos estudantes em conteúdos específicos (Ackerman,

1996a, 1996b, 2000), mais voltados a aplicação prática do conhecimento (situações práticas). Tais análises revelam que pela prova pretende-se avaliar a aplicação de conhecimentos em áreas específicas e que essa capacidade geral (1ª dimensão) do ponto de vista psicológico demonstra a necessidade de uma ênfase mais conteudista e assim cristalizada (Ackerman, 1996a, 1996b; Ackerman, 2005; Ackerman, Beier & Bowen, 2002; Castorina & Faigenbaum, 2002; Cronbach, 1957; McGrew, 2004, 2005) na diferenciação dos estudantes mais competentes.

Porém, apesar de na mesma prova existir a tendência em maior aparecimento de questões básicas de teorias administrativas e complexas, voltadas para a prática, o que assim também poderia polarizar os estudantes ingressantes e concluintes avaliados, outro elemento a ser discutido e assim levantado pelos juízes diz respeito ao nível de interpretação de textos dos estudantes e de suas habilidades de inferência por meio de tabelas e gráficos.

No próximo tópico dos resultados serão apresentadas informações mais qualitativas a respeito dos conteúdos e dificuldades dos itens apresentados no Mapa de Itens, assim como possíveis interpretações em termos de níveis de competência.

7. ESTUDO 2: Compreender o sentido dos níveis de proficiência empregando métodos da Teoria de Resposta ao Item

7.1 Método

Participantes

Participaram ativamente dessa análise 1 coordenador de curso e 4 professores, sendo que 2 dos professores relataram ter mais habilidade para avaliar as questões específicas de sua área de docência/ atuação. Cabe ressaltar, que tanto entre os juízes que aceitaram participar quanto aqueles que justificaram, uma questão relevante que foi levantada é que tal análise se torna bastante difícil visto que o perfil do docente e assim do curso é muito heterogêneo, envolvendo diferentes especificidades de conteúdo e perspectivas quanto à exigência e formação. Assim, na visão desses profissionais, é difícil para um único juiz avaliar todas as questões de forma qualitativa/ pedagógica não sendo de seu domínio particular ou maior. Portanto, foi sugerido que em outras pesquisas tal trabalho fosse realizado somente com coordenadores e/ou de forma grupal, trabalho que para esse trabalho não foi aceito/ ou possível nas instituições convidadas devido a outras questões institucionais levantadas.

Materiais

As análises dos juízes partiram do Mapa de Itens oriundo das análises da TRI, contendo informações sobre a dificuldade dos itens e a habilidade dos respondentes no ENADE 2006 de Administração.

Procedimento

Foram marcadas reuniões com os coordenadores de curso das instituições contatadas e enviados aos juízes as instruções de como executar as análises qualitativo-pedagógicas, como melhor descrito nos respectivos resultados. Foram convidados a participar como juízes para estabelecimento dos pontos de corte do Componente Específico da prova e assim interpretação dos resultados segundo os níveis de proficiência da escala, 4 coordenadores de curso de diferentes instituições de ensino e 16 professores/ pesquisadores que atuam ativamente e reconhecidamente na área.

Nesse procedimento os itens são organizados empiricamente em ordem crescente de dificuldade, estimada pela TRI, e apresentados aos juízes. A tarefa desses juízes era de determinar linhas de corte/ faixas de níveis de desempenho para classificação dos estudantes tendo como base os conhecimentos e habilidades necessários para a formação e atuação profissional.

Como se trata de um procedimento ainda pouco aplicado nas pesquisas e estudos educacionais, e mais especificamente nas instituições e pesquisadores convocados para participar, foi necessário a criação de uma espécie de “manual” para a explicação do mapa de itens e tarefa a ser realizada. Também a pesquisadora se dispôs e marcou reuniões para maior detalhamento da tarefa. Porém, devido ao tempo que deveria ser demandado para sua realização e o entendimento psicométrico mínimo necessário, muitos dos participantes acabaram por desistir e justificar a não participação nesse momento do estudo.

7.2 Resultados e Discussão

Nessa parte dos resultados foi feita uma análise da prova levando-se em consideração as especificações ligadas às competências, habilidades e conteúdos de cada

item da prova, de acordo com as diretrizes da prova (Anexo 2) elaboradas pela comissão assessora. Por intermédio do Modelo de Rasch foi criado um Mapa de Itens com a devida hierarquia de dificuldades e habilidades necessárias para que os juízes fizessem a análise de conteúdo desse mapa a fim de caracterizar os itens nos níveis de proficiência criados. Nesse tipo de procedimento é assumido que o nível de dificuldade dos itens reflete o nível de complexidade do conteúdo (Cizek, 1996; Cizek & Bunch, 2007; Impara & Plake, 2000; MacCann, & Stanley, 2006).

No presente estudo foi preciso que, a partir do mapa de itens do CE do ENADE 2006 de Administração, fossem realizadas a análise dos itens pelos juízes e indicados dois pontos na escala de habilidade que refletissem três níveis de proficiência. O primeiro deveria ser indicativo de qual agrupamento de questões os juízes avaliam que os alunos concluintes devem saber para serem considerados “minimamente competentes” ao final do curso. O segundo ponto chamado de “alta competência” deve referir a um nível de competência mínimo para se definir um aluno altamente competente, desejado para um aluno ao final do curso.

Diante dessa tarefa, os juízes avaliaram quais conteúdos, habilidades e competências refletem cada item ou conjunto de itens demarcados; se existem discrepâncias na classificação dos itens do ponto de vista estatístico (fáceis, médios ou difíceis) e qualitativo (buscar explicações em termos, por exemplo, dos distratores ou formato das questões); se alguns itens ou conjunto de itens refletem conhecimentos e estratégias mais próprias do início (ingressantes) ou final do curso (concluintes); entre outras informações que consideraram relevantes.

Seguindo esse procedimento foi feita uma análise exploratória do mapa de itens procurando identificar pontos significativos associados a blocos de questões ao longo da

escala e que pudessem definir níveis diferenciados de competência. Nessa análise procurou-se definir pontos que demonstrassem mudanças significativas em termos conteúdos/habilidades/competências implicadas nas questões.

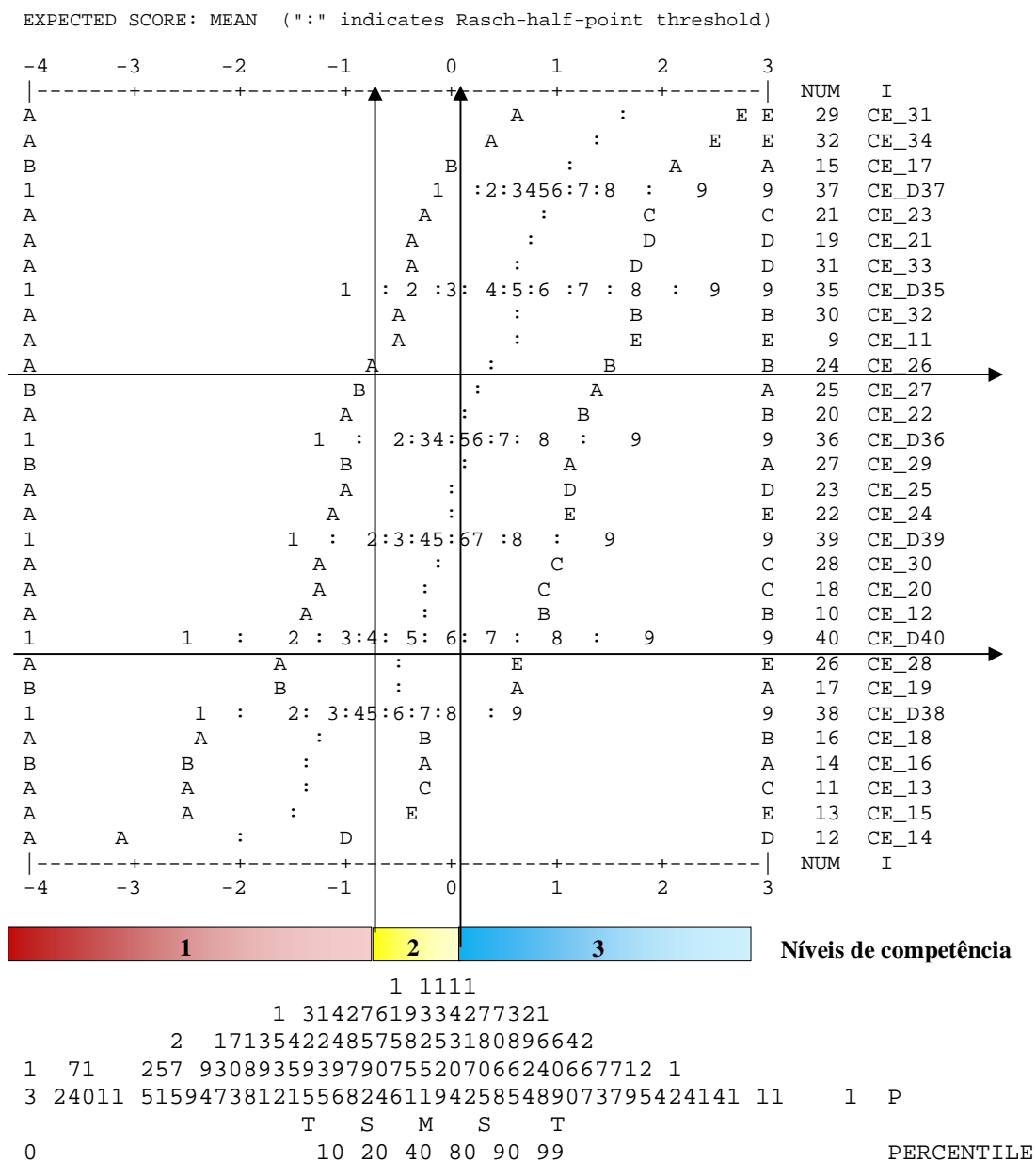


Figura 5 - Mapa dos itens do Componente Específico do ENADE 2006 de Administração com critérios

Como já apontado anteriormente, com a calibração do modelo por meio da TRI cada estudante recebe uma medida (theta) indicando sua posição na variável latente e cada item um parâmetro de dificuldade. Na medida em que o nível de competência aumenta em relação à dificuldade do item, suas chances de acertá-lo serão maiores do que 50%.

Nas regulamentações do SINAES há a proposição de uma escala de cinco níveis de competência, portanto seria interessante se fosse possível atender simultaneamente o critério de sentido substancial e interpretabilidade (Wolfe, 2007). Entretanto, é preciso ressaltar, conforme relatado anteriormente, que o índice de separação indicou que prova tem precisão para distinguir de um a dois estratos significativamente diferentes.

A partir da análise qualitativa e levando-se em conta os critérios anteriores foram definidos pelos juízes dois pontos, -0,5 e 0,15 (theta) dividindo a escala em três faixas limitadas pelos pontos de corte (insuficiente, minimamente competente e altamente competente). Com isso é possível verificar quais questões são mais prováveis de serem respondidas corretamente à medida que se progride no nível de competência implicado pelas faixas.

Faixa 1

Nesse nível o estudante tende a acertar os itens objetivos CE14, CE15, CE13, CE16, CE18, CE19 e CE28 (estes dois últimos estão no limite). A análise dessas questões mostra que elas requerem mais a habilidade de compreensão de textos, porém os juízes apontaram que as alternativas podem se mostrar bastante confusas, especialmente nas questões 15, 18 e 19. Em termos de conteúdo essas questões fazem exigência a um conhecimento sobre Teorias da Administração e de Recursos Humanos, geralmente mais ensinados até o 2º ou 3º ano do curso, com exceção do item 28 que exige conhecimentos de Sistemas e Tecnologia da Informação. Nas questões dissertativas, o estudante tem desempenho esperado de 5 ou 6 na CE_D38, e nas demais apresenta um padrão de no máximo 2, ou seja, insuficiente. Assim sendo esse nível sugere um nível de competência insuficiente ao se considerar o concluinte.

Faixa 2	<p>Nesse nível de competência o estudante tende a acertar, além dos itens da faixa anterior, os itens CE12, CE20, CE30, CE24, CE25, CE 29, CE22 e CE27 e a ter notas de 5-6 nas questões CE_D40, CE_D39 e CE_D36. Nesse nível de competência as questões requerem conhecimentos mais específicos sobre: Administração Mercadológica; Administração Financeira e Orçamentária, Contabilidade e Matemática; Administração de Serviços; Administração de Sistemas de Informação e Tecnologia de Informação; e Administração de Produção, Estatística e Matemática. Nesses itens o estudante precisaria mobilizar os conhecimentos básicos adquiridos nessas áreas para saber interpretar o que era perguntado e distinguir qual alternativa de resposta era correta, além de utilização de estratégias de resolução de problemas organizacionais, tomada de decisões e gestão de serviços, especialmente nas questões dissertativas. Por outro lado, nas questões dissertativas, esse nível de competência corresponde a um desempenho pouco mais satisfatório. Logo, a resolução das questões indicadas nesse nível requer a mobilização de conhecimentos específicos mais trabalhados a partir do meio do curso e é necessário um conhecimento mais aprofundado em termos de abrangência e contexto de aplicação, a fim de articular teoria e prática profissional. Assim sendo, indica um de nível competência mínimo.</p>
Faixa 3	<p>Nesse nível de competência, além dos itens das faixas anteriores, o estudante tende a acertar os itens CE11, CE32, CE33, CE21, CE23, CE17, CE 34 e CE31, porém com desempenho mediano nas questões dissertativas (acima de 4 na questão CE_D35 e acima de 2 na CE_D37). Esse é o nível mais alto de competência em que é esperado que os estudantes acertem as questões mais complexas e ainda mais voltadas para a prática profissional/ resolução de problemas, com exceção dos itens CE11 e CE17 cujos conteúdos são menos exigentes porém se mostram confusos. Além desses itens, os demais desse nível envolvem conhecimentos de Administração Mercadológica e Economia; Administração Estratégica; Administração Financeira e Orçamentária; Administração de Produção, Organização e Processos; e Administração de Recursos Materiais, Patrimoniais e Logística. Assim nesse nível o estudante tem que mobilizar conhecimentos específicos para interpretar dados e raciocinar sobre a coerência entre os dados e diferentes hipóteses explicativas. Esse nível também sugere a capacidade para compreensão do todo administrativo, dimensionamento de riscos, interação criativa, hierarquização de atividades e outros elementos críticos para interpretação da realidade das organizações. Deste modo sugere-se que esse nível de competência seja qualificado como nível de competência superior mínimo esperado para os concluintes.</p>

Outra informação relevante a apresentar se refere à relação existente entre a habilidade da pessoa, a dificuldade dos itens e a probabilidade de acertos podem ser expressas graficamente por meio da *Curva Característica dos Itens*. Nessas curvas,

apresentadas por item, é expresso no eixo X o nível de habilidade das pessoas e no eixo y a probabilidade de acerto ao item, sendo que a curva (linha cinza) expressa a probabilidade de acerto de pessoas com os vários níveis de habilidade tendo em vistas a dificuldade daquele item.

Esses gráficos apresentam a resposta correta marcada na legenda ao lado direito com dois asteriscos. Por essas informações é possível visualizar quais alternativas as pessoas com diferentes níveis de habilidade marcam em itens difíceis, médios e fáceis conforme os exemplos das Figuras a seguir (no Anexo 4 são apresentadas as curvas de todos os itens da prova). A primeira delas (Figura 6) é um dos itens referenciado como de competência mínima para o estudante concluinte, mas que segundo os juízes as categorias de respostas e mesmo o formato da questão podem ser considerados como elementos dificultadores para sua resposta correta. Esse comentário também é corroborado pelos juízes em itens semelhantes, como exemplo o item CE27, cuja exigência é categorizada como fácil/média, e que na opinião dos juízes teria mais acerto se fosse dissertativa ou mesmo em um formato diferente.

QUESTÃO 19

“A Iluminada” é uma empresa que produz luminárias. Ela atende a clientes individuais, vendendo desde luminárias para classes populares até produtos de luxo. A respeito das ofertas da empresa ao mercado, analise as afirmativas abaixo. Ao adotar segmentação de mercado, a empresa pode fixar preços diferenciados de acordo com os segmentos considerados.

PORQUE

As características dos clientes em cada segmento de mercado afetam o valor que eles percebem para os produtos da empresa.

Analisando as afirmações acima, conclui-se que

- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.

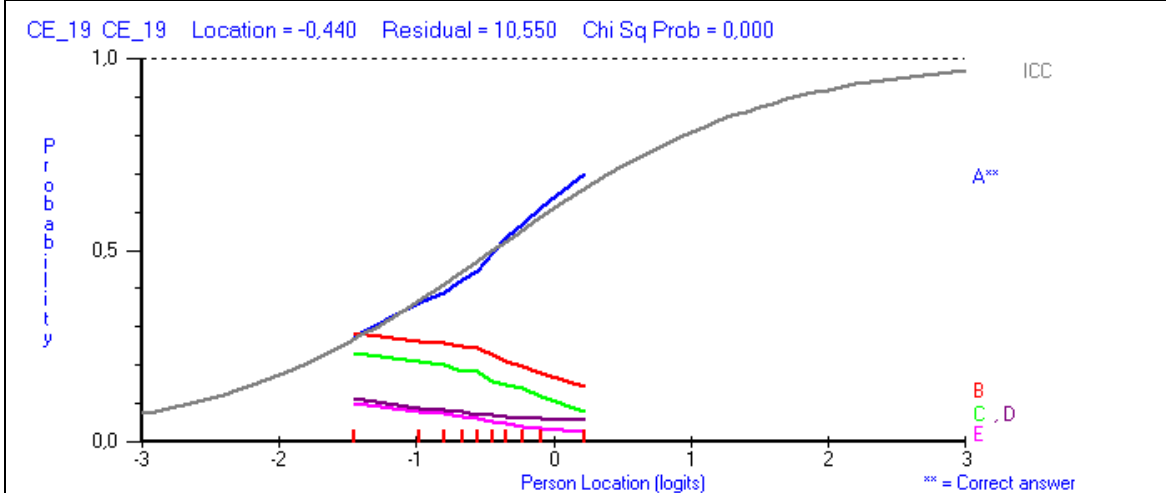


Figura 6. Relação entre habilidade e probabilidade de resposta por categoria no item CE19

Analisando a Figura 6 percebe-se que pessoas com níveis de habilidade de aproximadamente -1 optam também por responder nas demais categorias. Conforme a habilidade das pessoas aumenta, a resposta correta (A) passa a ser a mais provável para um nível de habilidade de até aproximadamente 0,0 em que a probabilidade é de 0,70 de acerto, mas em geral 22% dos respondem optam pela alternativa B.

Outro item apontado pelos juízes é o CE_12 em que apesar de mensurar de forma mais elementar as Teorias de Administração e ser um conteúdo ensinado no início do curso, a alternativa correta é assinada somente por 40% dos respondentes (Figura 7).

QUESTÃO 12

Embora constitua área de conhecimento das mais fascinantes, as bases teóricas da Administração ainda estão em formação. Os estudos pioneiros de Taylor e Fayol, por exemplo, foram ampliados, de forma significativa, nos anos posteriores. Sobre as Teorias da Administração pode-se afirmar que:

- I - na Burocracia, o trabalho realiza-se por meio de funcionários que ocupam cargos, os quais têm atribuições oficiais, fixas e ordenadas por meio de regras, leis ou disposições regimentais;
- II - na Administração Científica, enfatiza-se o estudo das tarefas, a seleção e o treinamento de trabalhadores e a busca pela eficiência operacional;
- III - na Reengenharia de Processos, há um esforço deliberado de se ter uma visão sistêmica da empresa, lastreado em estruturas organizacionais verticalizadas;
- IV - na visão Contingencial, procura-se analisar como as condições ambientais da empresa afetam as possibilidades de escolha nas decisões organizacionais;
- V - na abordagem Comportamentalista, a eficácia organizacional é promovida pela aplicação de análise quantitativa aos problemas e decisões administrativas.

São corretas, apenas, as afirmativas

(A) I, II e III.

(B) I, II e IV.

(C) I, III e V.

(D) II, IV e V.

(E) III, IV e V.

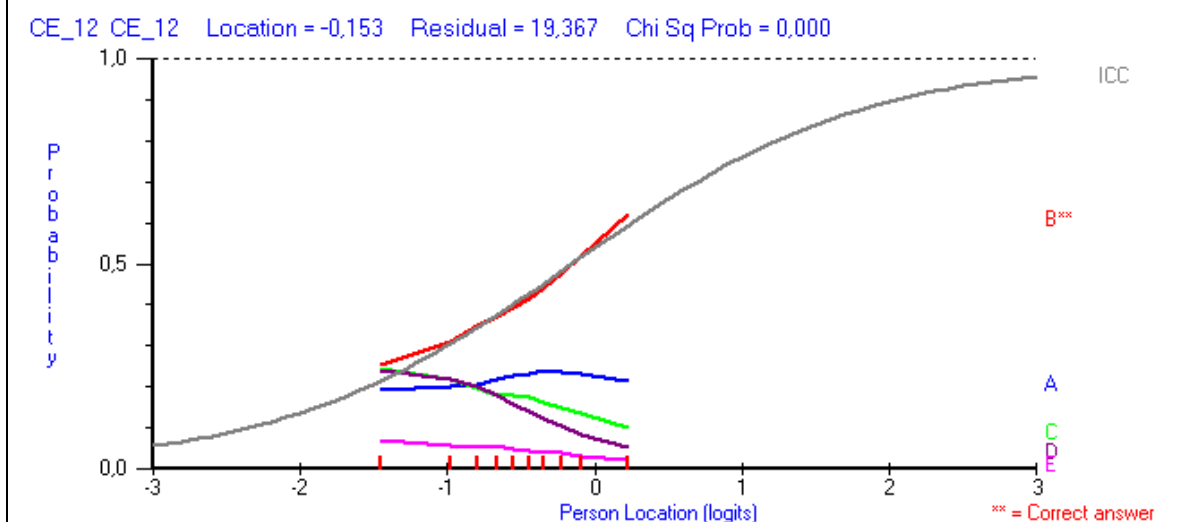


Figura 7. Relação entre habilidade e probabilidade de resposta por categoria no item CE12

Também foi mencionado que, na segunda faixa (competência mínima), além de conteúdos e habilidades passarem a ser mais voltados para a prática, alguns itens envolvem a interpretação de informações dispostas em gráficos e tabelas, como pode ser exemplificado pelo item CE30, em que menos de 40% dos estudantes assinalam a alternativa correta.

QUESTÃO 30

A Cia. Alonso de Auto Peças Ltda. distribui peças para oficinas de reparo de automóveis localizadas em grande área metropolitana. Embora se trate de um mercado competitivo, a Cia. Alonso gostaria de oferecer níveis de estoque adequados às oficinas atendidas, ao mesmo tempo em que deseja maximizar seus lucros. Ela é sabedora de que, à medida que aumenta a percentagem média de atendimentos aos clientes (nível de serviço), maior é seu custo de estoques. A fim de determinar a influência dos níveis de estoque no percentual de atendimento aos clientes, a Alonso fez um levantamento dos principais itens de seu estoque nos últimos seis meses. A seguinte tabela foi preparada:

Percentagem média de atendimento aos clientes	Nível médio mensal de estoque (R\$)	Custos de estoques mensais (R\$)	Receita média de vendas mensais (R\$)
80%	27.500,00	550,00	900,00
85%	30.000,00	600,00	1.200,00
90%	35.000,00	700,00	1.400,00
95%	40.000,00	800,00	1.450,00
98%	50.000,00	1.000,00	1.600,00

A partir dos dados apresentados nessa tabela, pode-se concluir que o maior lucro ocorrerá quando o nível de serviço for equivalente a

- (A) 80% (B) 85% (C) 90% (D) 95% (E) 98%

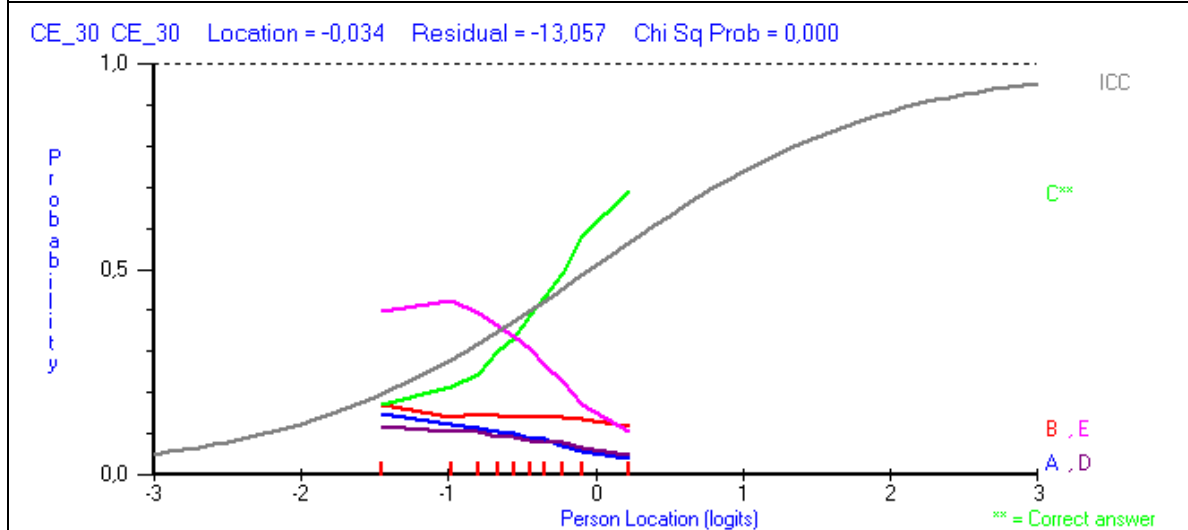


Figura 8. Relação entre habilidade e probabilidade de resposta por categoria no item CE30

Verifica-se também pela Figura 8 que até o nível de habilidade 0,5 a alternativa mais respondida é a E, considerada incorreta, sendo que a tendência é a diminuição dessas escolhas à medida que aumenta a capacidade dos estudantes, já que a alternativa correta passaria a ser mais escolhida entre os estudantes com alta competência. Outro item

apontado pelos juízes como sendo um conteúdo simples sobre tomada de decisão, porém cujo formato da questão se mostra confuso e pouco objetivo é o CE20 e o CE23.

QUESTÃO 23

O novo presidente da Empresa Mercados Múltiplos está redirecionando a sua estratégia por meio de um processo clássico centralizado de gestão estratégica de natureza prescritiva, cujos modelos conceituais e etapas são bem conhecidos.

Assim, o presidente está correto ao afirmar que o processo deve contemplar

- (A) o atendimento das expectativas dos diferentes grupos na estrutura de poder da empresa.
- (B) a implementação de estratégias emergentes.
- (C) a avaliação de concorrentes potenciais e produtos substitutos.
- (D) a utilização de processo participativo de tomada de decisão estratégica.
- (E) a manutenção do organograma e do sistema de informação existentes.

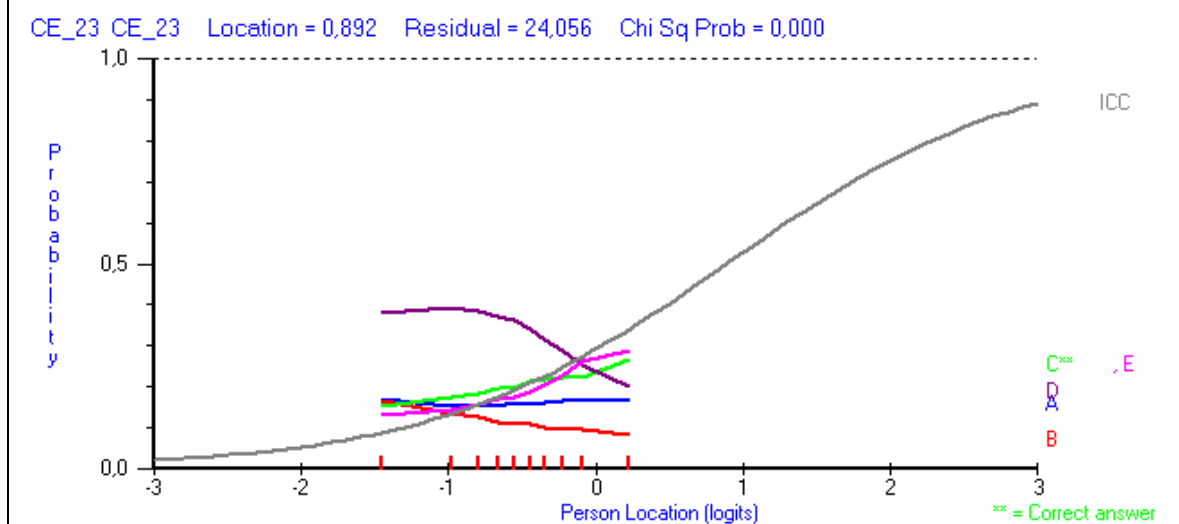
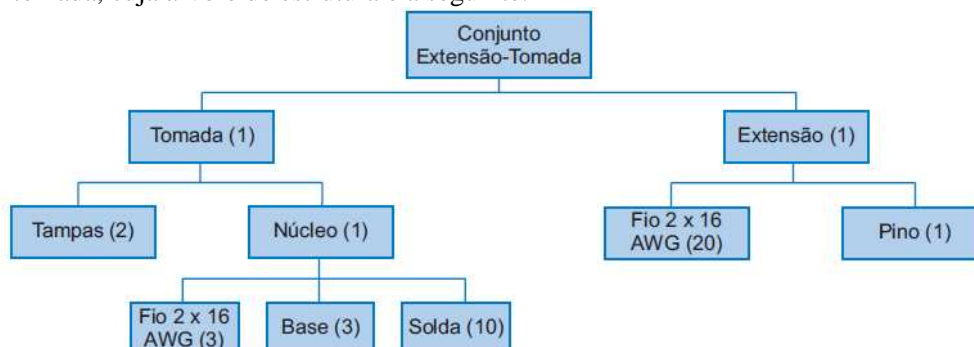


Figura 9. Relação entre habilidade e probabilidade de resposta por categoria no CE23

Ao analisar a questão 23 (Figura 9) conjuntamente com os juízes verifica-se que ocorreu provavelmente um problema na formulação do item. Os estudantes mais competentes não acertam mais do que os estudantes menos competentes, talvez por se confundirem com alternativas que também parecem possíveis, de forma que somente 20% das respostas estão assinaladas corretamente. Por fim, cabe ilustrar um item cujo conteúdo/habilidades/competências são altamente complexos (Figura 10).

QUESTÃO 34

A Cia. Goiás Velho S.A., fabricante de conectores, recebeu uma encomenda de 1.200 conjuntos extensão-tomada, cuja árvore de estrutura é a seguinte:



Os números entre parênteses referem-se às quantidades utilizadas na produção de cada conjunto. A Goiás Velho possui em estoque: extensão-tomada = 200; tomada = 100; extensão = 500; fio = 2.000. A nova política de estoques da empresa é a de não manter saldos em estoque, quer em conjuntos, quer em componentes.

A partir das informações apresentadas, pode-se concluir que a quantidade do componente fio (especificação 2 x 16 AWG) que precisa ser adquirido para atender a encomenda de 1.200 conjuntos extensão-tomada (utilizando todo o estoque existente) é

- (A) 25.600 (B) 21.000 (C) 12.700
(D) 11.000 (E) 10.700

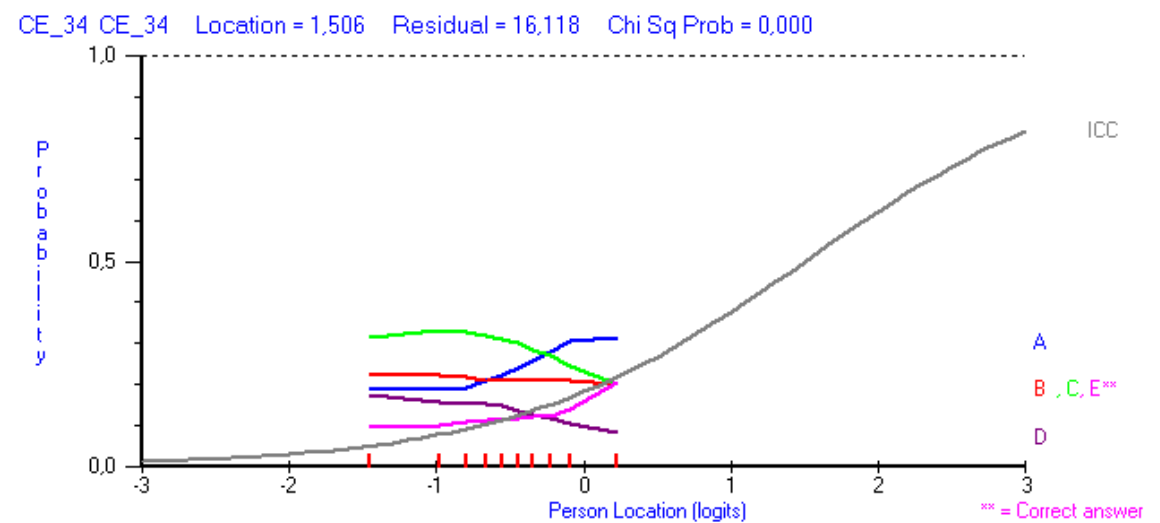


Figura 10. Relação entre habilidade e probabilidade de resposta por categoria no CE34

Nesse item, de acordo com as diretrizes, além da operação com valores e formulações quantitativas, e da articulação do conhecimento para a ação profissional, é exigido do estudante a compreensão dos processos produtivos e controle estatístico dos processos. Tais questões estão diretamente associadas à articulação do conhecimento para a prática profissional e são conteúdos exigidos quase que essencialmente no final do curso.

Entende-se que análises como essas permitem, em adição à interpretação referenciada a norma, atualmente usada no ENADE, a implementação de interpretações referenciadas ao critério (Smith, 2007), que se mostra mais adequada ao propósito do ENADE por permitir inferências e comparações que vão além da adequação do desempenho dos estudantes (Bond, 1996); e então permitindo maiores interpretações correspondentes às pontuações atribuídas (Embretson, 2006). Em termos práticos, também se pode futuramente utilizar essas informações para implementar melhorias nos cursos, identificando, por exemplo, falhas no processo de ensino e/ou avaliação ao entender mais amplamente os conhecimentos e habilidades cognitivas necessárias para a compreensão de determinadas temáticas (Rodrigues, 1994).

A partir desses resultados mais qualitativos, convém entender a configuração dos resultados dos estudantes nessa nova escala criterial. Do total de participantes do exame, 56% estão classificados no nível de competência insuficiente, 38% na faixa minimamente competente e somente 5% como altamente competente. Esses resultados podem demonstrar o rigor dos juízes e também o panorama dos estudantes do Brasil nesse ano do ENADE, a partir dessas classificações. A Tabela 11 mostra a distribuição de estudantes nos critérios segundo o momento do curso.

Tabela 11- *Distribuição dos estudantes segundo critérios e momento do curso*

Critério	Estatísticas	Momento no curso	
		Ingressante	Concluinte
Insuficiente	N	47.873	25.503
	%	65,2	34,8
Competência Mínima	N	21.546	28.229
	%	43,3	56,7
Alta Competência	N	2.077	4.334
	%	32,4	67,6

Verifica-se que com o aumento das faixas, aumenta também a representação de estudantes concluintes, de forma que na Alta competência esses totalizam 67% dos participantes. Também esses estudantes foram divididos de acordo com o conceito atribuído ao curso (Tabela 12).

Tabela 12 - *Distribuição de estudantes no CE por faixa de competência, momento e conceito do curso*

Momento no curso	Conceito	Estatísticas segundo medida de critério							
		Insuficiente		Competência Mínima		Alta Competência		Total	
		N	%	N	%	N	%	N	%
Ingressantes	Conceito 1	746	83,9	133	15,0	10	1,1	889	100,0
	Conceito 2	13.176	75,9	3.935	22,7	239	1,4	17.350	100,0
	Conceito 3	21.659	67,1	9.784	30,3	817	2,5	32.260	100,0
	Conceito 4	3.539	53,5	2.792	42,2	290	4,4	6.621	100,0
	Conceito 5	725	27,5	1.478	56,2	429	16,3	2.632	100,0
	Total	39.845	66,7	18.122	30,3	1.785	3,0	59.752	100,0
Concluintes	Conceito 1	513	74,0	168	24,2	12	1,7	693	100,0
	Conceito 2	9.203	58,8	5.990	38,3	453	2,9	15.646	100,0
	Conceito 3	13.269	42,3	16.142	51,5	1.942	6,2	31.353	100,0
	Conceito 4	1.460	23,6	3.822	61,8	905	14,6	6.187	100,0
	Conceito 5	268	10,3	1.384	53,4	939	36,2	2.591	100,0
	Total	24.713	43,8	27.506	48,7	4.251	7,5	56.470	100,0

Nota-se pela Tabela 12 uma diminuição de estudantes de acordo com o aumento dos conceitos na faixa de competência insuficiente. Ao analisar os resultados dos ingressantes classificados como Minimamente Competentes, grande parte está com conceito 3. E, dos estudantes ingressantes com conceito 5 no ENADE, 16% já se apresentam na classificação de Alta competência, faixa adequada a um estudante concluinte, conforme os pontos de corte estabelecidos pelos especialistas. No caso dos concluintes, a amostra apresenta também uma concentração maior de estudantes com conceito 3, e, dos estudantes com

conceito 5 no ENADE, 36% se apresentam na classificação de Alta competência. A Tabela 13 apresenta a comparação segundo categoria administrativa.

Tabela 13 - *Distribuição de estudantes no CE por faixa de competência, momento e categoria administrativa*

Momento no curso	Categoria administrativa	Estatísticas segundo medida de critério							
		Insuficiente		Competência Mínima		Alta Competência		Total	
		N	%	N	%	N	%	N	%
Ingressantes	Federal	718	32,1	1.220	54,5	300	13,4	2.238	100,0
	Estadual	1.094	49,5	961	43,5	155	7,0	2.210	100,0
	Municipal	1.383	73,0	478	25,2	33	1,7	1.894	100,0
	Particular	44.678	68,6	18.887	29,0	1.589	2,4	65.154	100,0
	Total	47.873	67,0	21.546	30,1	2.077	2,9	71.496	100,0
Concluintes	Federal	405	16,8	1.367	56,8	634	26,4	2.406	100,0
	Estadual	619	28,9	1.219	56,9	306	14,3	2.144	100,0
	Municipal	748	42,0	914	51,3	118	6,6	1.780	100,0
	Particular	23.731	45,9	24.729	47,8	3.276	6,3	51.736	100,0
	Total	25.503	43,9	28.229	48,6	4.334	7,5	58.066	100,0

Tanto para ingressantes quanto para os concluintes, a maior parte dos estudantes provém de instituições de ensino particulares. Dentre os ingressantes das federais 54,5% dos estudantes estão localizados na faixa de competência mínima, enquanto que nesse mesmo momento do curso nos estabelecimentos municipais e particulares prevalece a ocorrência de estudantes com competência insuficiente. Para os concluintes, mostra-se o predomínio dos estudantes considerados minimamente competentes e entre os respondentes das federais 26% se apresentam no nível de alta competência. A seguir, indica-se a distribuição de estudantes de acordo com o tipo de organização acadêmica, em que prevalecem estudantes universitários e de Faculdades, Escolas e Instituições Superiores (Tabela 14).

Tabela 14 - Distribuição de estudantes no CE por faixa de competência, momento e organização acadêmica

Momento no curso	Organização acadêmica	Estatísticas segundo medida de critério							
		Insuficiente		Competência Mínima		Alta Competência		Total	
		N	%	N	%	N	%	N	%
Ingress.	Universidade	12.647	60,6	7.330	35,1	907	4,3	20.884	100,0
	C. Universitário	5.624	66,7	2.588	30,7	216	2,6	8.428	100,0
	Fac. Integradas	2.941	71,7	1.078	26,3	80	2,0	4.099	100,0
	Fac., Esc. e Inst. S.	26.653	70,1	10.522	27,7	869	2,3	38.044	100,0
	C. de Educ. Tecn.	8	19,5	28	68,3	5	12,2	41	100,0
	Total	47.873	67,0	21.546	30,1	2.077	2,9	71.496	100,0
Concl.	Universidade	6.437	36,7	9.128	52,0	1.983	11,3	17.548	100,0
	C. Universitário	3.341	43,2	3.855	49,8	544	7,0	7.740	100,0
	Fac. Integradas	1.796	48,6	1.733	46,9	164	4,4	3.693	100,0
	Fac., Esc. e Inst. S.	13.917	48,0	13.465	46,4	1.620	5,6	29.002	100,0
	C. de Educ. Tecn.	12	14,5	48	57,8	23	27,7	83	100,0
	Total	25.503	43,9	28.229	48,6	4.334	7,5	58.066	100,0

Dentre os ingressantes, mais de 60% dos estudantes encontram-se na faixa insuficiente de competência, com exceção dos Centros de Educação Tecnológica, cuja maioria se localiza no grupo de minimamente competentes. Ainda, mais de 45% dos estudantes se mostram, ao final do curso, em situação de competência mínima, de forma que as formas de organização acadêmica apresentam um aumento de respondentes considerados minimamente competentes, em comparação aos ingressantes. E, nos Centros Tecnológicos, houve um aumento na faixa de alta competência entre os concluintes.

Na Tabela 15 são apresentadas as estatísticas referentes à análise multivariada testando a significância estatística das diferenças entre as médias de habilidade dos estudantes (theta) dividindo-se o grupo dois momentos (ingressante e concluinte) e em grupos de estudantes de cursos com conceitos diferenciados, bem como a interação entre essas duas variáveis.

Tabela 15 - *Análise de variância fatorial múltipla para testar diferenças de habilidade entre os estudantes segundo o conceito e o momento do curso*

Fonte	Soma de quadrados	gl	Média dos Quadrados	Estatística <i>F</i>	Sig.	Eta Parcial ao quadrado
Modelo	64368,042	1	64368,042	209685,795	0,000	0,643
Momento do curso	474,745	1	474,745	1546,534	0,000	0,013
Conceito	3163,173	4	790,793	2576,094	0,000	0,081
Momento * conceito	139,870	4	34,967	113,910	0,000	0,004
Erro	35674,038	116212	,307			
Total	301394,000	116222				

Verifica-se que há efeito significativo devido à interação entre momento do curso e conceito, o que indica que parte da variância das medidas de theta está sistematicamente associada às características dessas variáveis. Na última etapa dos resultados do presente estudo serão testadas as variações de desempenho dos estudantes após controle de algumas variáveis possivelmente mais explicativas das diferenças entre as instituições.

8. ESTUDO 3: Investigar as evidências de validade por intermédio da verificação das relações com outras variáveis

A fim de verificar a validade do ENADE para a avaliação da qualidade do curso poderia ser testada a relação do desempenho do estudante com outros indicadores de qualidade disponíveis no próprio curso e instituição. Porém, é necessário primeiramente verificar qual a variabilidade útil do exame para medir o efeito-escola e então, posteriormente, realizar tais análises correlacionais (quando a variabilidade encontrada for vantajosa).

8.1 MÉTODO

Participantes

Foram considerados para a análise somente os estudantes concluintes e que haviam respondido as duas partes da prova, o que totalizou em 58.066 estudantes de 1.121 instituições do país, participantes do ENADE 2006.

Materiais

Banco de dados contendo as informações sobre o desempenho dos estudantes, perfil socioeconômico e características das instituições participantes do ENADE 2006 de Administração.

Procedimento

Os dados foram organizados a fim de incluir nas análises as variáveis mais significativas disponíveis do perfil prévio dos estudantes (provenientes do questionário

socioeconômico) para analisar o efeito dos cursos e instituições no desempenho dos estudantes, como será visto de forma mais pormenorizada nos resultados.

8.2 Resultados e Discussão

Nessa etapa, foi usada a Análise de Regressão Hierárquica ou os Modelos Multiníveis (Raudenbush & Bryk, 2002; Snijders & Bosker, 2003) a partir do Programa Computacional MLwiN em sua versão 2.18. O modelo multinível ou hierárquico respeita a estrutura de agrupamento da população em estudo de uma forma parcimoniosa e fornece a 'ordenação' intrínseca das instituições condicionada às variáveis de controle fundamentais em cada área de estudo (Ferrão, Leite & Beltrão, 2001).

Primeiramente foram eliminados da análise os estudantes que responderam menos de 23 questões a fim de selecionar apenas aqueles que responderam 75% dos itens do Componente Específico. E, utilizou-se o resultado do ENADE dos estudantes concluintes no CE convertido na escala theta como a variável dependente de interesse.

Tabela 16 - Estatísticas dos estudantes no CE por momento do curso

Momento	Estatísticas	
Ingressantes	Número de estudantes	71.496
	Média	-0,74
	Desvio Padrão	0,54
	Mínimo	-4,87
	Máximo	3,02
Concluintes	Número de estudantes	58.066
	Média	-0,46
	Desvio Padrão	0,48
	Mínimo	-4,84
	Máximo	2,26

Essa distribuição do desempenho dos concluintes pode ser mais bem visualizada na Figura 11, de forma que cruzando com as informações agregadas pela análise de juízes, verifica-se que os concluintes estão em média localizados na faixa de competência insuficiente (N=25.503) e competência mínima (N=28.229).

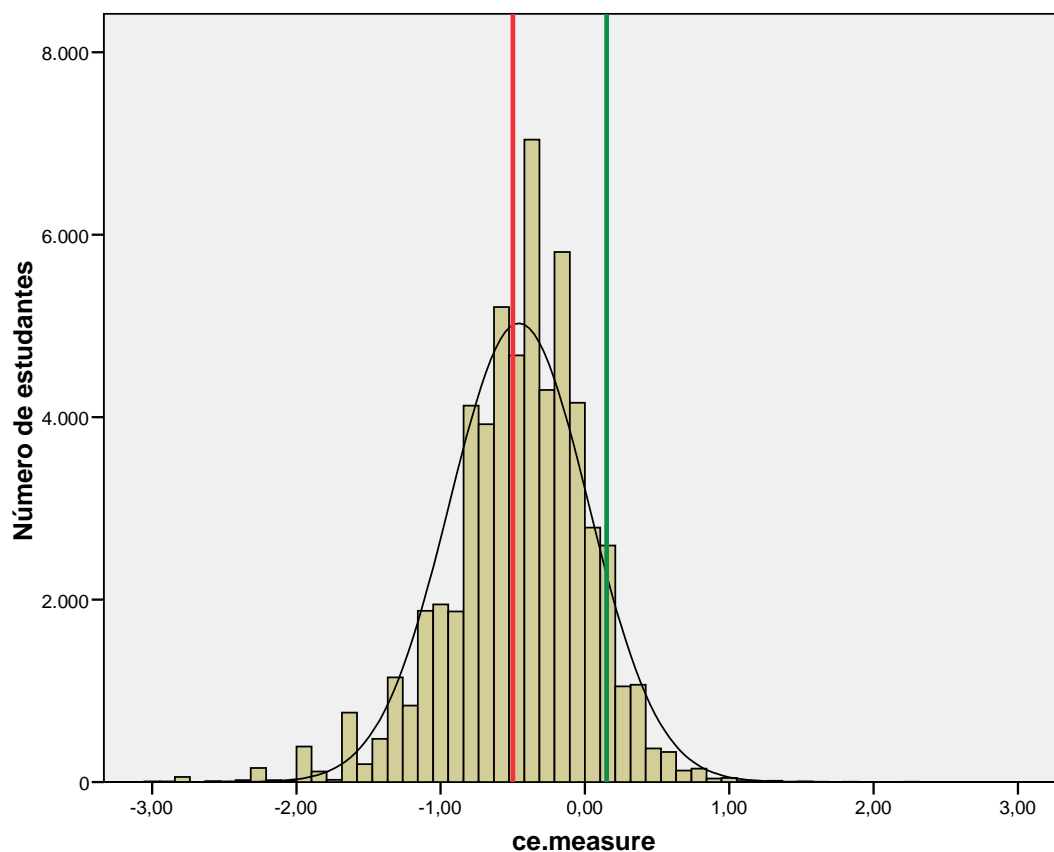


Figura 11. Distribuição de habilidades no componente específico (concluintes)

Também é necessário atentar-se para as fontes de variação devidas ao nível socioeconômico dos alunos individualmente (nível 1), pois para revelar a verdadeira contribuição da escola, precisa-se antes controlar o impacto das condições socioeconômicas (Jesus & Laros, 2004). Como fatores explicativos do perfil prévio dos estudantes no Nível 1 foram selecionadas do questionário socioeconômico algumas variáveis que possam ser mais relevantes para esta explicação.

Especificamente em relação às variáveis socioeconômicas do questionário do ENADE, Silva e cols. (2010) apresentaram os resultados da análise desse baseado no ENADE de 2005 por meio da criação de variáveis que congregam grupos de itens que se esperavam estar associados ao desempenho do estudante no exame. Esses autores partiram da análise fatorial do questionário, das correlações dos itens com o desempenho no componente específico da prova, além da análise qualitativa das questões a fim de agrupá-las e assim utilizá-las de forma mais sintetizada em estudos sobre a influência das instituições na qualidade da educação.

As variáveis criadas foram: a) Domínio de línguas, que agrega itens que avaliam a percepção do aluno do seu domínio de língua inglesa e espanhola; b) Hábitos de Leitura, variáveis que descrevem os hábitos de leitura cotidianos, como tempo dedicado a leitura e estudo, assuntos mais lidos e uso de bibliotecas; c) Atividades Acadêmicas Extracurso, itens que descrevem as atividades extracurriculares realizadas pelo aluno que estão associadas à vida acadêmica e à participação em atividades culturais; d) Uso do computador, que congrega os itens que abordam a frequência, o local e a finalidade do uso do computador; e) Acervo e serviços da Biblioteca, com itens que avaliam a disponibilidade de serviços, equipamentos e materiais bibliográficos na biblioteca; f) Condições e recursos da instituição, itens em que os alunos avaliam as instalações físicas, equipamentos, materiais de consumo e acessos aos recursos da instituição, e da adequação dessas instalações, equipamentos, materiais e acesso a recursos às necessidades das aulas práticas; g) Formação social, reflexão de problemas sociais, que agrega itens que avaliam a percepção do aluno da contribuição do curso para sua compreensão de problemas sociais; h) Formação social e política, itens que avaliam a percepção que o aluno tem da contribuição do curso na sua capacidade de compreender temas gerais do cotidiano e

aspectos sociais, político e culturais da realidade brasileira; i) Avaliações de aspectos pedagógicos e de ensino, itens que avaliam os planos, os procedimentos de ensino; j) Oferecimento e contribuição da Iniciação Científica, itens em que o aluno cita se participou de projetos de iniciação científica e qual a relevância dessa participação em sua formação; l) Oferecimento e contribuição da Monitoria, itens em que o aluno cita se participou como monitor, e qual a relevância dessa participação em sua formação; m) Oferecimento e contribuição de Projetos de Extensão, itens em que o aluno cita se participou de projetos de extensão e qual a relevância dessa participação em sua formação; n) Desenvolvimento de Competências Metodológicas, itens que avaliam a percepção dos alunos da contribuição do curso na sua capacidade de utilizar recursos informatizados, procedimentos metodológicos, analisar dados e assimilar criticamente conteúdos; o) Competências Sociais, que agregam itens que avaliam a percepção dos alunos da contribuição do curso na sua capacidade de atuar em equipe, comunicar informações, tomar decisões e solucionar problemas, agir eticamente e raciocinar de maneira crítica; e p) Competências, que agrega as variáveis Competências metodológicas e Competências Sociais.

Além dessas agregações, alguns itens se mostraram importantes a serem analisados separadamente, o que reduziu as 103 perguntas do questionário a 37 questões/agrupamentos, a fim de facilitar as análises. Dessas questões, foram selecionadas no nível 1 quatro que se mostraram mais associadas com o desempenho na prova e também indicadas na literatura (ex. Ferrão, 2003), como pode ser verificado na Tabela 17.

Assim sendo, as variáveis explicativas analisadas nos modelos construídos compreendem itens obtidos a partir dos questionários respondidos pelos alunos e alguns itens obtidos por agregação, de acordo com o estudo de Silva e cols. (2010). A Tabela 15 descreve as diferentes variáveis inseridas nos modelos estatísticos construídos. A variável

‘*Tipo de Instituição*’ não foi inserida na análise por não ser o objetivo desse estudo e a variável ‘*Escolaridade dos pais*’ foi inserida durante o processo de modelagem, mas descartada em etapas posteriores por não mais se mostrar com um efeito significativo.

Tabela 17- *Escalas das variáveis explicativas do modelo de regressão multinível*

Nível	Variável	Descrição	Codificação
Nível 1	Tipo_instit	Tipo de escola	1-Pública 2-Mista 3-Privada
	escol_pais	Escolaridade dos pais	Variando de 1 a 5 (nenhuma escolaridade a Ensino Superior)
	renda	Renda familiar	Variando de 1 a 7 (Até 3 salários-mínimos a Mais de 30 salários-mínimos)
	Sus_trab	Formas de sustento e trabalho	Variando de 1 a 5 (não trabalha a principal responsável pelo sustento)
Nível 2	ce.measure_mean.0	Pontuação no CE na escala theta de habilidade (ingressantes)	-3,0 a 3,0
	nt_fg_mean.0	Desempenho na FG (ingressantes)	0 a 100

As variáveis independentes foram centralizadas a fim de permitir comparação. Primeiramente foi analisado o modelo nulo, sem variáveis explicativas, para identificar qual parte da variância total pode ser atribuída aos níveis 1 e 2. A notação formal do modelo é a seguinte:

$$\begin{aligned}
 y_{ij} &= \beta_{0j} + e_{ij} \\
 \beta_{0j} &= \beta_0 + \mu_{0j} \\
 \mu_{0j} &\sim N(0, \sigma_{\mu_j}^2) \\
 e_{ij} &\sim N(0, \sigma_e^2)
 \end{aligned}$$

Onde, y_{ij} representa o desempenho na parte específica do ENADE do estudante i no curso j . Esse desempenho é previsto por um parâmetro β_{0j} composto pela média geral da amostra β_0 e pela média dos concluintes do curso j de qual o estudante i é parte μ_{0j} que representa o desvio da IES em relação à média geral e mais um resíduo e_j relacionado ao estudante indicando quanto ele se afasta da previsão baseada na média da IES. Assim esse modelo separa a variância total em dois componentes: (a) a variância dos estudantes ao redor da média de seu curso σ_e^2 pertencente ao Nível 1 e (b) a variância da média dos estudantes de uma IES j ao redor da média geral $\sigma_{\mu_j}^2$ pertencente ao Nível 2.

Vale salientar que o modelo tenta explicar a variância total ($\sigma_{\mu_j}^2 + \sigma_e^2$) empregando os parâmetros da IES no Nível 2 e assim serve para explicada $\sigma_{\mu_j}^2$ versus a variância não explicada σ_e^2 permitindo-se calcular o coeficiente de correlação intraclasse (r_{ICC}) dividindo a variância do nível macro (IES) sobre a variância total e que portanto pode variar de 0 a 1 dada pela fórmula:

$$r_{ICC} = \frac{\sigma_{\mu_j}^2}{\sigma_{\mu_j}^2 + \sigma_e^2}$$

Quando seu valor é nulo ou se aproxima de zero, isso significa que as instituições são homogêneas entre si e que o desempenho escolar do aluno é independente da escola que frequenta (Ferrão & Fernandes, 2003; Jesus & Laros, 2004), e a análise multinível não precisa ser aplicada. No caso do presente estudo, tal resultado indicaria que a diferença nos desempenhos dos estudantes não pode ser atribuída necessariamente a IES frequentada.

Portanto nesse modelo no Nível 1 (N = 45.550) estão os estudantes concluintes agrupados dentro das instituições de ensino do Brasil que participaram do ENADE 2006 do curso de Administração que, por sua vez, e constituem o Nível 2 (N = 1.121). De acordo

com o modelo nulo a estimativa da variância entre as instituições é 0,031, a variância entre os estudantes dentro das instituições é igual a 0,204 e a variância total 0,235. Deste modo, pôde-se calcular o coeficiente de correlação intraclasse que mede o quanto da variação total decorre das diferenças entre as instituições. Segundo o resultado observado somente 13,2% da variação total se deve à variação entre as instituições.

$$\begin{aligned}
 & \text{ce.measure}_{ij} \sim N(XB, \Omega) \\
 & \text{ce.measure}_{ij} = \beta_{0ij} \text{cons} \\
 & \beta_{0ij} = -0.467(0.006) + u_{0j} + e_{0ij} \\
 & [u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.031(0.002)] \\
 & [e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [0.204(0.001)] \\
 & -2 * \loglikelihood(IGLS Deviance) = 72969.064(56699 \text{ of } 56699 \text{ cases in use})
 \end{aligned} \tag{1}$$

Também por intermédio da Figura 12 que mostra os valores do parâmetro μ_{0j} das 1.121 instituições ordenadas pela magnitude, do menor ao maior, é possível visualizar que existe pouca variabilidade entre as instituições, de forma a inferir que há um pequeno efeito das instituições no desempenho dos estudantes, ainda sem a utilização de nenhuma variável de controle.

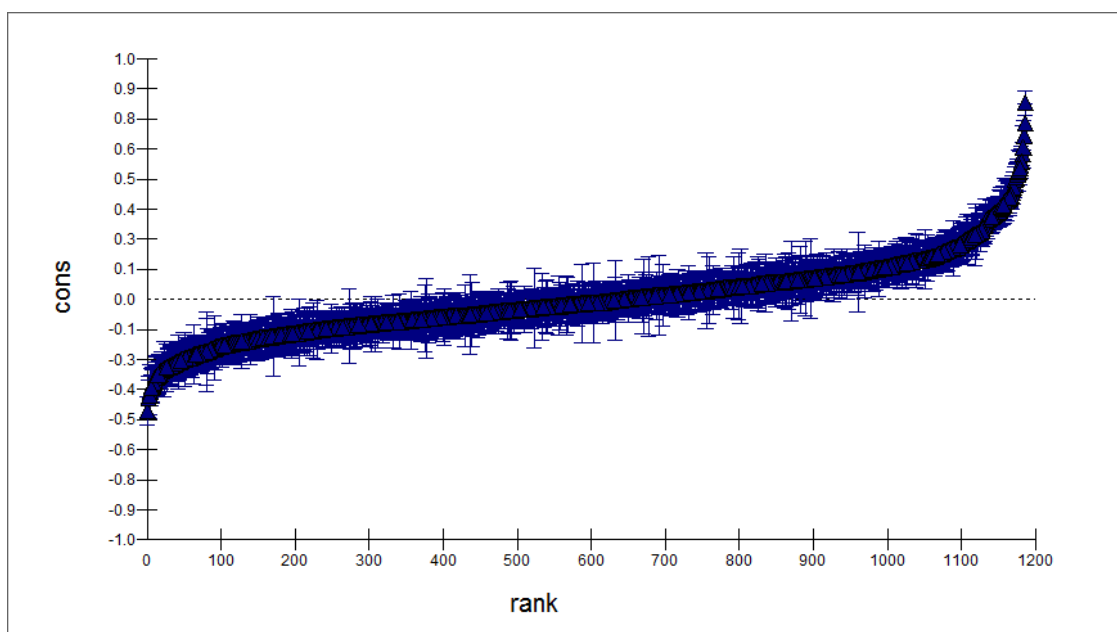


Figura 12. Ranking das instituições no Componente Específico no modelo nulo

Verifica-se por essa figura que, sem controle, a maior parte das instituições tem estudantes classificados como ‘mínima competência’ e somente um pequeno número como ‘alta competência’ (“calda” do gráfico).

Em seguida foram testados modelos incluindo-se variáveis externas tanto no Nível 1 (nível socioeconômico do aluno, por exemplo) quanto no Nível 2 (média no ENADE dos estudantes ingressantes da instituição) com o objetivo de remover a variância irrelevante à avaliação instituição. A notação geral do modelo é:

$$\begin{aligned}
 y_{ij} &= \beta_{0j} + (X_{1ij} + \dots + X_{Vij}) + e_{ij} \\
 \beta_{0j} &= \beta_0 + (\beta_{1j} + \dots + \beta_{Vj}) + \mu_{0j} \\
 \mu_{0j} &= \sim N(0, \sigma_{\mu_j}^2) \\
 e_{ij} &= \sim N(0, \sigma_e^2)
 \end{aligned}$$

No modelo os termos $X_{1ij} + \dots + X_{Vij}$ indicam as V variáveis X_{Vij} predictoras sendo consideradas no Nível 1. Já $\beta_{1j} + \dots + \beta_{Vj}$ indicam as V variáveis predictoras β_{Vj} do Nível 2. Nota-se que, no Nível 1, somente se consideraram efeitos fixos, isto é, não se testou se

estes efeitos variavam entre IES's já que isso complicaria o modelo e fugiria do objetivo buscado nessa parte do trabalho. É importante ressaltar que a variância total dos termos nesse modelo ($\sigma_{\mu_j}^2 + \sigma_e^2$) será menor que o modelo nulo, e a magnitude da redução indica o montante de variância explicado pelas variáveis que foram inseridas:

$$r = \frac{\sigma_{Tot_nulo}^2 - \sigma_{Tot_mod1}^2}{\sigma_{Tot_nulo}^2}$$

Como já foi referido, o Modelo 1 começa incluindo, a priori, as principais variáveis de Nível 2 relacionadas ao perfil cognitivo dos estudantes que buscam a IES, a saber: Pontuação no CE na escala theta de habilidade (ingressantes) e desempenho na FG (ingressantes) como medida de controle inicial. A utilização da média de habilidade dos estudantes ingressantes no modelo ocorre pois não se tem uma medida individual do desempenho do concluinte quando ele entrou na IES (os dados não são longitudinais) assim uma aproximação dessa medida seria o desempenho médio dos ingressantes que é usado como um preditor do desempenho dos concluintes. Portanto se faz esse controle no Nível 2. O resíduo dessa análise representa uma parcela de variância mais “pura” pois remove parte da variância que já se sabe não estar associada à qualidade da IES. Os resultados do modelo estão apresentados a seguir.

$$\begin{aligned} ce.measure_{ij} &= \beta_{0j} + 0.524(0.035)(ce.measure_mean.0-gm)_{ij} + 0.008(0.001)(nt_fg_mean.0-gm)_{ij} + e_{ij} \\ \beta_{0j} &= -0.459(0.004) + u_{0j} \\ u_{0j} &\sim N(0, \sigma_{u0}^2) \quad \sigma_{u0}^2 = 0.012(0.001) \\ e_{ij} &\sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0.206(0.001) \\ -2*loglikelihood &= 72586.225(56643 \text{ of } 58428 \text{ cases in use}) \end{aligned} \quad (2)$$

Verificou-se que, do total de variação do desempenho dos concluintes no Componente Específico do ENADE somente 5% pode ser atribuído as instituições ($\sigma_{\mu_j}^2$) após o controle do perfil anterior (de entrada) dos estudantes ainda que feito no Nível 2.

Em seguida, incluiu-se no modelo as variáveis preditoras do Nível 1: Escolaridade dos pais, Renda familiar, Formas de sustento e trabalho. Como já citado, foram incluídas outras variáveis mas mantidas somente as que tiveram efeito significativo. O modelo com a adição dessas variáveis reduziu a 4% a variância potencialmente relevante atribuída às instituições ($\sigma_{\mu_j}^2$), como segue:

$$\begin{aligned}
 \text{ce.measure}_{ij} &= \beta_{0j} + 0.504(0.035)(\text{ce.measure_mean.0-gm})_{ij} + \\
 &\quad 0.009(0.001)(\text{nt_fg_mean.0-gm})_{ij} + 0.029(0.002)(\text{renda-gm})_{ij} + \\
 &\quad 0.055(0.002)(\text{Sus_trab-gm})_{ij} + e_{ij} \\
 \beta_{0j} &= -0.453(0.004) + u_{0j} \\
 u_{0j} &\sim N(0, \sigma_{u0}^2) \quad \sigma_{u0}^2 = 0.011(0.001) \\
 e_{ij} &\sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0.201(0.001) \\
 -2*\text{loglikelihood} &= 57251.288(45510 \text{ of } 58428 \text{ cases in use})
 \end{aligned}
 \tag{3}$$

Ao visualizar a Figura 13, semelhante à Figura 12, mas refeita considerando o último modelo, é possível verificar que a variabilidade entre as instituições é ainda menor no modelo final, de forma a inferir que as diferenças se dão mais entre os alunos dentro das instituições e não entre instituições. Entre as melhores instituições a diferença média é de 0,1 logit e entre as piores, 0,4. E, apesar da divisão em cinco faixas de conceitos, proposta pelo INEP verifica-se que, controlando-se algumas variáveis irrelevantes, o montante de variância útil é muito pequeno

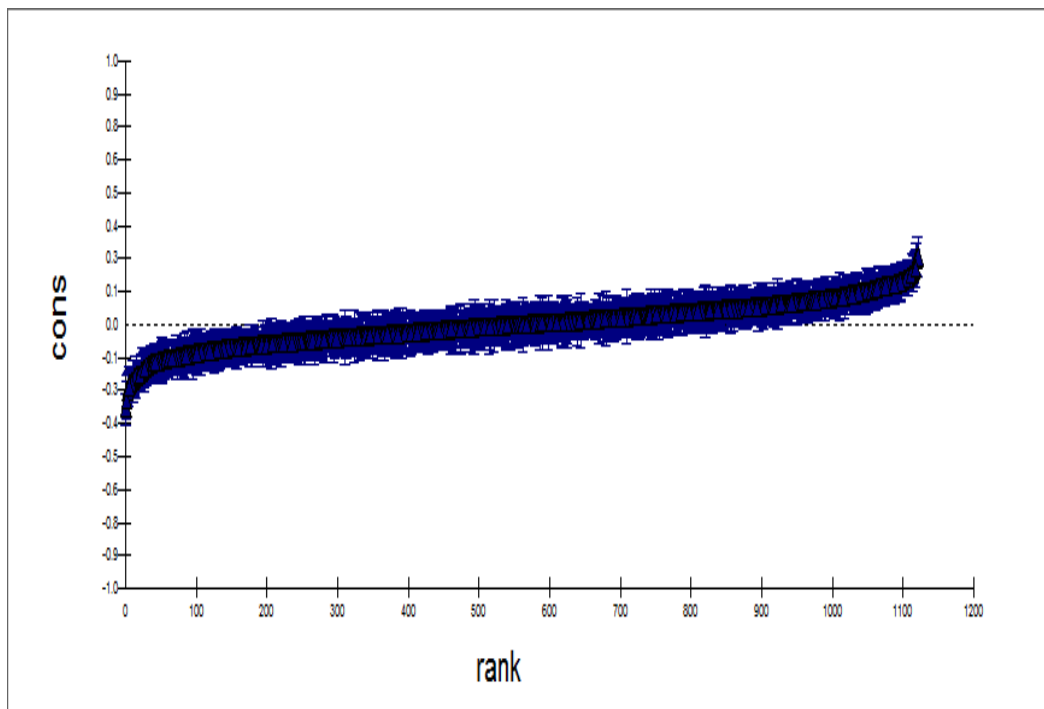


Figura 13. Ranking das instituições no Componente Específico após controle

Após controle, verifica-se que não há quase nenhuma variância em termos de níveis de competência, de forma que a melhor IES é em média 0,5 logit melhor que a pior. É possível analisar os efeitos das variáveis explicativas de ambos os níveis e os resultados encontrados podem ser vistos na Tabela 18, que apresenta as estatísticas referentes ao modelo nulo e final encontrado nesse estudo.

Tabela 18 - Modelo sem variáveis explicativas e após a inserção das variáveis de controle

Variáveis explicativas	Modelo Nulo (sem variáveis explicativas)			Modelo de Referência (após inserção de variáveis de controle)		
	Efeito	EP	Razão T	Efeito	EP	Razão T
Efeitos Fixos						
Intercepto (β_0)	-0,467			-0,453		
ce.measure_mean.0 (β_{1j})				0,504	0,035	14,4
nt_fg_mean.0 (β_{2j})				0,009	0,001	9
Renda (X_{1ij})				0,029	0,002	14,5
Sus_trab (X_{2ij})				0,055	0,002	27,5
Efeitos Randômicos						
Nível 2 (variância de $\mu_{0j}, \sigma_{\mu j}^2$)	0,031			0,011		
Nível 1 (variância de $\theta_{ij}, \sigma_{\theta}^2$)	0,204			0,201		
Correlação intra-classe		0,13			0,04	
Deviance		72969,064			57251,288	
Número de Parâmetros estimados		3			7	
Diferença nos Deviances					3929,444	
Diferença do número de parâmetros					4	

Cabe destaque ao valor do intercepto (-0,467), que representa o valor médio da proficiência dos estudantes concluintes e que nesse estudo está expresso em termos de habilidade (theta), cuja escala assume valores entre -3,0 e 3,0. Verifica-se que o desempenho dos estudantes concluintes é consideravelmente menor que o esperado nas faixas de competência estabelecidas anteriormente e mais próximo do valor esperado para os estudantes intermediários no curso.

Os resultados sugerem que a maior parte da variância do desempenho médio entre as instituições no ENADE devem-se a diferenças na composição dos seus estudantes, o que indica um efeito muito grande do perfil inicial dos mesmos nas instituições. Tais dados apontam a relevância de se controlar estatisticamente o efeito dessas variáveis, caso contrário se atribuiria aos cursos e assim também às instituições um efeito que não cabe a elas (Belline, 1996; Brooke & Soares, 2008; Castro, 2001; Ferrão, 2003; Ferrão &

Fernandes, 2003; Flouri, Buchanan & Bream, 2002; Opdenakker & Van Damme, 2000; Opdenakker & Van Damme, 2005, Primi, 2006; Primi & Nunes, no prelo).

Pode-se verificar também que a correlação intraclasse diminuiu para 0,04 depois da inserção das variáveis de controle. Este valor é compatível ou até menor ao valor da correlação intraclasse de outros estudos (Andrade & Laros, 2007; Ferrão & Fernandes, 2003; Jesus & Laros, 2004; Soares, 2004; Soares & cols., 2004), apesar da ressalva de que esses não foram voltados para o Ensino Superior brasileiro.

Com essas análises pretendeu-se verificar a proporção de variância relevante disponível no desempenho dos estudantes e, ao mesmo tempo, investigar fatores associados a esse componente que poderiam ser considerados explanatórios da variabilidade da qualidade entre os cursos. Assim, conseguiu-se decompor a variância do desempenho do estudante concluinte no ENADE separando-se a parcela mais relevante na construção de indicadores de qualidade dos cursos.

Deste modo, as diferenças encontradas estão mais centradas no perfil dos estudantes dentro das instituições e mais associada ao conhecimento anterior dos estudantes (Soares, Ribeiro & Castro, 2001). Pode-se concluir a importância da profundidade dos conhecimentos adquiridos no percurso acadêmico como fator diferenciador na competência dos estudantes (Ackerman, 1996a; Horn & Cattell 1966; McGrew, & Flanagan, 1998; Primi e cols. (2001) e, assim, a importância da promoção de experiências adequadas de aprendizagem ao longo da vida escolar e avaliação dessas de forma geral (Ackerman, 1996a, 1996b; Ackerman, 2000; Ackerman 2005; Ackerman, Beier & Bowen, 2002; Castorina & Faigenbaum, 2002; Cronbach, 1957; Hambrick, 2005; Smith & Marsiske, 1994).

Enquanto implicação prática, essas questões vão ao encontro dos postulados de vários autores que destacam a dificuldade de usar dados transversais para descrever a contribuição específica da escola no desempenho dos alunos (Oliveira, Franco & Soares, 2007; Primi, 2006; Primi & cols., 2009b; Rutter & cols., 2008a; Willms, 2008); a importância de elementos para a interpretação dos resultados educacionais (Bond, 1996; Embretson, 2006; Ferraz & cols., 1994; Rodrigues, 1994; Tamassia, 2005) e de uma medida precisa de controle da influência da aprendizagem anterior (Gray, 2008; Mortimore & cols., 2008a, 2008b; Rutter & cols., 2008a, 2008b; Soares, Ribeiro & Castro, 2001).

Uma última etapa foi analisar diferenças de desempenho dos estudantes das instituições segundo os conceitos no ENADE sem controle das variáveis dos estudantes e após o controle, modelo final da análise multinível (Tabela 19).

Tabela 19 - *Diferenças de desempenho dos estudantes por conceito antes e após controle de variáveis externas*

Conceito	Antes controle		Após controle	
	Média	Erro padrão da média	Média	Erro padrão da média
Conceito 1	-0,84	0,017	-0,05	0,019
Conceito 2	-0,63	0,004	-0,02	0,004
Conceito 3	-0,45	0,003	0,00	0,002
Conceito 4	-0,24	0,006	0,02	0,006
Conceito 5	0,02	0,009	0,04	0,010

Antes de controlar as variáveis externas no modelo multinível, a média geral do grupo segundo os conceitos era de -0,46, faixa de competência mínima. Os resultados indicaram que os estudantes de instituições com conceitos maiores tendem a possuir desempenhos mais elevados do que os demais ($F[4, 56465] = 1716,98; p < 0,05$), o que denota também a importância do curso e instituição de formação.

Depois do controle as diferenças entre conceitos quase com nenhuma variância, sendo que a média geral do grupo passou a ser 0,00, mesma faixa de competência ($F[4, 45399] = 16,55; p > 0,05$). Tais resultados demonstram, portanto, que os maiores conceitos estão intimamente relacionados a maiores níveis de competência, o que pode permitir a inferência de que os diferentes conceitos implicam diferenças importantes de níveis de habilidade no momento de entrada no curso. De forma geral, verificou-se que a habilidade e a qualidade de conteúdo e aprendizagem adquirida anteriormente à formação universitária têm uma influência direta na então quantidade e qualidade da formação cursada (Ackerman, 1996a, 1996b; Ackerman, 2005; Dongo, 2008; Falcão, 1985; Primi, 2008). E, também indica a importância e contribuição da Psicologia (mais especificamente das teorias de aprendizagem e inteligência) para o entendimento dessas medidas educacionais e nas possíveis explicações para as variações de desempenho entre estudantes e instituições.

9. CONSIDERAÇÕES FINAIS

O objetivo geral desse trabalho foi discutir a validade em avaliar a qualidade dos cursos a partir de medidas do desempenho dos estudantes, como proposto no ENADE. Primeiramente verificou-se o predomínio unidimensional que foi inferido como significando que a prova, em geral, avalia o domínio de conhecimentos específicos ($Gc/Gkn/gk$) condizente com a literatura psicológica que aponta tais domínios como principais diferenciadores na comparação entre os estudantes mais competentes. Porém, muitas das questões da prova analisadas podem ser resolvidas com a habilidade adequada em leitura e escrita, entendida como Grw no modelo CHC. Esse achado permitiu verificar a contribuição da psicologia na explicação do desempenho dos estudantes avaliados pelo ENADE, de forma que apesar do exame no Componente Específico priorizar a avaliação dos conhecimentos específicos da área (Gkn), na prática outras habilidades como Gf e Grw demonstram ser importantes na resolução dessas questões específicas e imprescindíveis para o alcance da alta competência. Tal discussão demonstra a relevância dos achados de autores como Ackerman (2000) que apontam a necessidade de considerar as medidas processuais (gp) tanto quanto de conhecimento (gk) na mensuração cognitiva em adultos.

Apesar do ajuste do modelo aos dados de acordo com a TRI, especialmente no Componente Específico, entende-se que seja essencial melhorar a prova, especialmente em sua extensão (quantidade de itens X conteúdos/habilidades/competências que se almeja avaliar) e formato (deveria incluir mais questões dissertativas para abarcar o que está previsto nas diretrizes). Entende-se que a maneira como o exame está formatado não é suficiente para diferenciar os estudantes em termos de competência.

Além desses pontos, ao longo do trabalho foi discutida a necessidade de que estudos ou indicadores que partem dos resultados do ENADE sejam feitos a partir de um

delineamento longitudinal, questão metodológica que não ocorre nesse exame que parte de dados transversais e também utiliza a mesma prova para avaliar simultaneamente ingressantes e concluintes. Tal ponto também mostra a necessidade de, por exemplo, se partir de estudos longitudinais que incluam medidas de controle do conhecimento inicial dos estudantes para assim possibilitar a inferência do efeito dos cursos e das instituições no desempenho dos respondentes.

A utilização da TRI para o exame dos itens juntamente com a análise realizada pelos juízes permitiu verificar de forma mais qualitativa e criterial os resultados encontrados e, principalmente possibilitou interpretar os escores de forma mais substancial. Tal procedimento de escalonamento comportamental usado pelos juízes apresenta-se vantajoso para outros estudos, pois ao ser aplicado às avaliações educacionais possibilita estabelecer que aspectos foram aprendidos pelo estudante ao longo do curso. Também devido às limitações das interpretações baseadas na norma, entende-se ser essencial a mudança do sistema de referência à norma para referência ao critério, a fim de proporcionar análises mais qualitativas e cujas comparações tenham como base critérios pré-estabelecidos para avaliar as aprendizagens individuais; além de, por consequência, desencorajar o uso dos rankings entre as instituições. Para tanto, é necessário que as provas tenham qualidade condizente com os pressupostos da TRI (unidimensionalidade essencial) e assim uma redefinição do domínio das provas. Considera-se ser essencial a realização de análises desse tipo visto que, por mais que a avaliação se dê a partir dos resultados do produto da aprendizagem, entende-se que esses podem levar a novos direcionamentos e planejamentos sobre o próprio processo de aprendizagem.

Por fim, a fim de verificar a variância relevante do desempenho do estudante que poderia carregar informações sobre a qualidade do curso, encontrou-se que somente 4% da

variabilidade se devem ao efeito dos cursos e instituições. Entende-se que essa variabilidade, já na literatura apontada como baixa, demonstra a inviabilidade dos indicadores de qualidade do curso a partir somente dos resultados do exame. Além disso, o Componente Específico também se apresenta pouco preciso e discriminativo, sendo esse um dos motivos pelo qual pouco da variância do ENADE consegue explicar efeito escola (uma parcela importante de sua variância é erro). Tais apontamentos são relevantes se considerado o impacto das avaliações como o ENADE em vários níveis (acadêmico, institucional, político, social, pessoal, entre outros), e, portanto, por possibilitar questionamentos sobre as limitações do exame e das interpretações dele decorrentes (vinculadas na mídia por meio da publicação das notas das instituições).

Uma das principais questões decorrentes desses últimos resultados é a constatação de que parte das diferenças entre os cursos é explicada pelos distintos desempenhos dos ingressantes e por isso não pode ser “atribuída” ao curso. E que, portanto, por maior que possa ser a qualidade dos cursos de algumas instituições, dificilmente essas atingirão níveis mais altos em razão do perfil dos alunos ingressantes. Deste modo, sugere-se que sejam criadas formas de correção da nota da instituição a partir do desempenho dos ingressantes, de maneira que sejam ponderadas as diferenças de competências desses estudantes no momento de entrada no curso. Tal forma de correção poderia ser realizada, por exemplo, por intermédio do cálculo do efeito-escola obtido pelos modelos multiníveis em que, inclusive, poderiam ser consideradas outras variáveis além daquelas pesquisadas no presente estudo.

Considera-se que de forma mais abrangente, o presente estudo leva a uma questão fundamental do processo avaliativo: repensar os padrões educacionais que ainda carregam um modelo de educação que visa à classificação dos alunos e os transforma em notas, sem

considerar toda a bagagem de conhecimento que é incorporada com as experiências de vida e durante anos de estudos. Assim sendo, verifica-se a contribuição essencial da psicologia para a área ao sugerir formas mais eficazes de avaliar, que vão além das sugestões tradicionais em educação

Os dados apresentados foram frutos de um processo analítico-interpretativo mas não encerra as possibilidades de diálogo com os dados obtidos e com a literatura abordada. Ao contrário, com esta pesquisa abre-se uma proposta de diálogo para outras investigações possíveis, como forma de contribuir para a reflexão e a discussão sobre temas relacionados ao processo avaliativo, formação e fatores de diferenciação dos estudantes nos vários níveis de ensino.

Entende-se que, apesar de no presente estudo ter sido analisado somente um curso e ano do exame, a abordagem metodológica utilizada pode ser aplicada aos demais cursos e anos avaliados pelo ENADE. Assim sendo, convém que sejam realizados mais estudos que confirmem tais achados e apontem outras questões psicométricas talvez mais específicas de cada área de conhecimento avaliada pelo exame e, por conseguinte, contribuam para a busca de evidências de validade do exame e de sua utilização no processo de avaliação do Ensino Superior. Também em estudos futuros é preciso o envolvimento de uma maior quantidade de juízes de regiões e instituições diferentes; além da inclusão de outras variáveis pessoais e de contexto na mensuração do efeito das instituições no desempenho dos estudantes.

10. REFERÊNCIAS

- Ackerman, P. L. (1987). Individual Differences in Skill Learning: An Integration of Psychometric and Information Processing Perspectives. *Psychological Bulletin*, 102(1), 3-27.
- Ackerman, P. L. (1996a). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, 22, 227-257.
- Ackerman, P. L. (1996b). Adult intelligence. *Practical Assessment, Research & Evaluation*, 5(8). Disponível em: <<http://pareonline.net/Articles.htm>>. Acessado em: 01/04/2008.
- Ackerman, P. L. (2000). Domain-Specific Knowledge as the “Dark Matter” of Adult Intelligence: Gf/Gc, Personality and Interest Correlates. *Journal of Gerontology: Psychological Sciences*, 55B(2), 69-84.
- Ackerman, P. L. & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: development of a battery for predicting air traffic controller success. *Journal of applied psychology*, 78(3), 413-432.
- Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences*, 33, 587–605.
- Ackerman, P. L., & Cianciolo, A. T. (2002). Ability and Task Constraint Determinants of Complex Task Performance. *Journal of Experimental Psychology: Applied*, 8(3), 194–208.
- Ackerman, P. L., & Beier, M. E. (2003). Trait Complexes, Cognitive Investment, and Domain Knowledge. Em R. J. Sternberg & E. L. Grigorenko. *Perspectives on the*

- psychology of abilities, competencies, and expertise* (pp. 1-30). New York: Cambridge University Press.
- Ackerman, P. L., & Beier, M. E. (2005). Knowledge and Intelligence. Em O. Wilhelm & R.W. Engle. *Handbook of understanding and measuring intelligence* (pp. 125-139). California: Sage Publications.
- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, 131(1), 30–60.
- Almeida, L.S. (1988). *Teorias da inteligência*. Porto: Edições Jornal de Psicologia.
- Almeida, L. S. (1994). *Inteligência: definição e medida*. Aveiro: Cidine.
- Almeida, L. S., Lemos, G., Guisande, M. A., & Primi, R. (2008). Inteligência, escolarização e idade: normas por idade ou série escolar? *Avaliação Psicológica*, 7(2), 117-125.
- Alves, I. C. B. (1998). Variáveis significativas na avaliação da inteligência. *Psicologia Escolar e Educacional*, 2(2), 109-114.
- Alves, M. T. G., & Franco, C. (2008). A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias*. (pp. 482-500). Belo Horizonte: Editora UFMG.
- Alves, M. T. G. & Soares, J. F. (2007a). As pesquisas sobre o efeito das escolas: contribuições metodológicas para a Sociologia da Educação. *Sociedade e Estado*, 22(2), 435-473.

- Alves, M. T. G. & Soares, J. F. (2007b). Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos. *Educação em Revista*, 45, 25-58
- Alves, M. T. G. & Soares, J. F. (2008). O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no Ensino Fundamental. *Educação e Pesquisa*, 34(3), 527-544.
- Alves, M. T. G. & Soares, J. F. (2009). Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. *Opinião Pública*, 15(1), 1-30.
- American Educational Research Association-AERA, American Psychological Association-APA, & National Council on Measurement in Education-NCTM (1999). *Standards of Educational and Psychological testing*. Washington, DC: American Educational Research Association.
- Anastasi, A., & Urbina, S. (2000). *Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Andrade, D. F. (2001). Comparando desempenhos de grupos de estudantes por intermédio da Teoria de Resposta ao Item. *Estudos em Avaliação Educacional*, 23(2), 31-69.
- Andrade, D. F., Tavares, H. R., & Valle, R. C. (2000). *Teoria da Resposta ao Item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística.
- Andrade, J. M., & Laros, J. A. (2007). Fatores associados ao desempenho escolar: estudo multinível com dados do SAEB/2001. *Psicologia Teoria e Pesquisa*, 23 (1), 33-41.
- Barreyro, G. B. (2004). Do Provão ao SINAES: o processo de construção de um novo modelo de avaliação da Educação Superior. *Avaliação Institucional*, 9(1), 37-49.

- Bartholomeu, D., Silva, M. C. R., Pires, S. D., & Primi, R. (2010). Estabelecimento de pontos de corte pela TRI para uma prova equalizada com o ENADE. Em F. C. Capovilla. (Org.). *Transtornos de aprendizagem: Progressos em avaliação e intervenção preventiva e remediativa* (p. 229-238), v. 1, São Paulo: Memnon.
- Belloni, I. (1996). A universidade e o compromisso da avaliação institucional na reconstrução do espaço social. *Avaliação, 1*(2), 5-14.
- Berg, C. A. (2000). Intellectual Development in Adulthood. Em R. J. Sternberg. *Handbook of Intelligence* (pp. 117-140). New York: Cambridge University Press.
- Bittencourt, H. R., Viali, L., Casartelli, A. O., & Rodrigues, A. C. M. (2008). Uma análise da relação entre os conceitos Enade e IDD. *Estudos em Avaliação Educacional, 19*(40), 247-262.
- Bond, L. (1996). Norm- and Criterion-Referenced Testing. *Practical Assessment, Research & Evaluation, 5*(2). Disponível em: <<http://ericae.net/pare/getvn.asp?v=5&n=2>>. Acesso em: 2 abr. 2009.
- Bowles, S., & Gintis, H. (2008). Reforma da escola em retrospectiva. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 90-104). Belo Horizonte: Editora UFMG.
- Brandão, Z. (2000). Fluxos Escolares e Efeitos Agregados pelas Escolas. *Em aberto, 17*(71), 41-48.
- Brito, M. R. F. (2006). *A avaliação dinâmica da aprendizagem em um exame de larga escala para os cursos de Engenharia de Produção*. [on line]. Disponível em: <http://www.abepro.org.br/arquivos/websites/12/GTG2_Marcia_Brito.pdf>. Acesso em: 23 fev. 2009.

- Brito, M. R. F. (2007). ENADE 2005: Perfil, desempenho e razão da opção dos estudantes pelas Licenciaturas. *Avaliação*, 12(3), 401-443.
- Brito, M. R., Munhoz, A., Primi, R., Gonçalves M. H., Rezi, V., Neves, L. F., Sanches, M. H., & Marinheiro, F. B. (2000). Exames nacionais: Uma análise do ENEM aplicado à Matemática. *Revista Avaliação*, 5(4), 445-54.
- Brooke, N., & Soares, J. F. (Org.). (2008). *Pesquisa em eficácia escolar: origem e trajetórias*. Belo Horizonte: Editora UFMG.
- Burlamaqui, M. G. B. (2008). Avaliação e Qualidade na Educação Superior: tendências na literatura e algumas implicações para o sistema de avaliação brasileiro. *Estudos em Avaliação Educacional*, 19(39), 133-154.
- Camargo, E. A. P. (2007). *Políticas Públicas de Educação Nacional e Paulista: A função social do Ensino Médio*. Dissertação de Mestrado. Universidade cidade de São Paulo.
- Carroll, J. B. (1993a). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (1993b). Test theory and the behavioral scaling of test performance. Em *Frederiksen, N., Mislevy, R. J. & Bejar, I. I. (Eds.). Test theory for a new generation of tests*. (pp. 297-322) Hillsdale, NJ: Lawrence Erlbaum.
- Carroll, J. B. (1997). The Three-Stratum Theory of Cognitive Abilities. Em D. P. Flanagan, J. L. Gensheft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 122-130). New York: The Guilford Press.
- Carvalho, L. F., Silva, M. C. R., Dias, A. S., Primi, R., Nunes, C. H. S. S., Vendramini, C. M. M., Miguel, F. K., Nogueira, S. O. (2007). Análise Psicométrica da prova Pluridisciplinar do curso de Enfermagem e estudo comparativo com os resultados do

- ENADE 2004. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- Castorina, J. A., & Faigenbaum, G. (2002). The epistemological meaning of constraints in the development of domain knowledge. *Theory & Psychology*, 12(3), 315-334. Disponível em: <<http://tap.sagepub.com/cgi/content/abstract/12/3/315>>. Acesso em: 16 ago. 2008.
- 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. (2001). O novo debate sobre o uso de padrões na educação. *Latin America Basic Education Summit Miami – Florida*. Disponível em: <http://www.inep.gov.br/download/internacional/idioma/O%20Novo%20Debate%20sobre%20o%20uso%20de%20Padr%C3%B5es%20na%20Educa%C3%A7%C3%A3o_Po rtugu%C3%AAs_Ingl%C3%AAs_Espanhol.pdf>. Acesso em: 3 ago. 2008.
- Cattell, P. (1940). *The measurement of intelligence of infants and young children*. New York: Psychological Corporation.
- Cattell, Raymond B. (1957). *Culture fair intelligence test, a measure of "g": Scale 3, forms A and B (high school pupils and adults of superior intelligence)*. Savoy, IL: Institute for Personality and Ability Testing.
- Cattell, R. B. (1971). *Intelligence: its structure, growth and action*. Boston: Houghton Mifflin.
- Cerri, L. F. (2004). Saberes históricos diante da avaliação do ensino: notas sobre os conteúdos de história nas provas do Exame Nacional do Ensino Médio – ENEM. *Revista Brasileira de História*, 24(48), 213-231.

- César, C. C., & Soares, J. F. (2001). Desigualdades acadêmicas induzidas pelo contexto escolar. *Revista Brasileira de Estudos de População*, 18(1/2), 97-110.
- Cizek, G. J. (1996). Standard-Setting Guidelines. *Educational Measurement Issues and Practices*, 15(1), 13-21
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: a guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Coleman, J. S. (2008). Desempenho nas escolas públicas. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias*. (pp. 26-32). Belo Horizonte: Editora UFMG.
- Colombi, F. C. C. (2004). A Repercussão do Exame Nacional do Ensino Médio na Escola Pública Estadual: do currículo a avaliação, sob a ótica de professores e alunos. *Dissertação de Mestrado*. Universidade do Vale do Itajaí
- Condé, F. N., & Laros, J. A. (2007). Unidimensionalidade e a propriedade de invariância das estimativas da habilidade pela TRI. *Avaliação Psicológica*, 6, 205-215.
- Connell, M. W., Sheridam, K., & Gardner, H. (2003). On Abilities and Domains. Em: R. J. Sternberg & E. L. Grigorenko. *The psychology of abilities, competencies, and expertise* (pp. 1- 30). New York: Cambridge University Press.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Cronbach, L. J. (1996). *Fundamentos da Testagem Psicológica*. Porto Alegre: Artes Médicas.
- Dias, A. S., Silva, M. C. R., Vendramini, C. M. M., & Brito, M. R. F. (2007). Hábitos de leitura dos estudantes de matemática que participaram do ENADE 2005. *III Congresso*

Brasileiro de Avaliação Psicológica e XII Conferência Internacional da Avaliação Psicológica: Formas e Contextos, João Pessoa.

Dias, A. S., Silva, M. C. R., Carvalho, L. F., Primi, R., Nunes, C. H. S. S., Vendramini, C. M. M., Miguel, F. K., Nogueira, S. O. (2007). *Análise Psicométrica da prova Pluridisciplinar do curso de Nutrição e estudo comparativo com os resultados do ENADE 2004*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.

Dongo M., A. (2008). La teoría del aprendizaje de Piaget y sus consecuencias para la praxis educativa. *Revista de Investigación em Psicologia* [Online], 11(1), 167-181. Disponível em: http://pepsic.bvs-psi.org.br/scielo.php?script=sci_arttext&pid=S1609-74752008000100011&lng=pt&nrm=iso. Acesso em: 22 abr. 2009.

Embretson, S. E. (2006). The Continued Search for Nonarbitrary Metrics in Psychology. *American Psychologist*, 61(1), 50-55.

Falcão, G. M. (1985). *Psicologia da aprendizagem*. 2ª edição. São Paulo: Ática.

Ferrão, M. E. (2003). *Introdução aos modelos de regressão multinível em Educação*. Campinas, São Paulo: Komedi.

Ferrão, M. E., & Beltrão, K. (2001). *Tracing schools which do not penalise over age students*. Trabalho apresentado na 27th Annual Conference of the International Association for Educational Assessment, Rio de Janeiro.

Ferrão, M. E., Beltrão, K., Fernandes, C., Santos, D., Suarez, M., & Andrade, A. (2001). O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos de População*, 18(1/2), 111-130.

- Ferrão, M. E., Leite, I. C. & Beltrão, K. I. (2001). *Introdução à modelagem multinível em avaliação educacional*. Rio de Janeiro: Escola Nacional de Ciências Estatísticas.
- Ferrão, M. E., Beltrão, K., & Santos, D. (2002). Políticas de não-repetência e a qualidade da educação: evidências obtidas a partir da modelagem dos dados da 4ª série do Saeb-99. *Estudos em Avaliação Educacional*, 26,47-73.
- Ferrão, M. E., & Fernandes, C. (2003). O Efeito-Escola e a Mudança – Dá para mudar? Evidências da investigação brasileira. *Revista Electrónica sobre Calidad, Eficácia y Cambio em Educación (REICE)*, 1(1) [on line]. Disponível em: <<http://www.ice.deusto.es/rinace/reice/vol1n1/FerraoyFernandes.pdf>>. Acesso em: 8 jun. 2009.
- Ferraz, M. J., Carvalho, A., Dantas, C., Cavaco, H., Barbosa, J., Tourais, L., & Neves, N. (1994). A Avaliação Criterial /Avaliação Normativa. I.I.E. *Pensar Avaliação, Melhorar a Aprendizagem*. Lisboa: Instituto de Inovação Educacional.
- Flanagan, D. P., & McGrew, K. S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: joint confirmatory factor analysis of the WJ-R and KAIT in a non-white sample. *Journal of School Psychology*, 36(2), 151-182.
- Flanagan, D. P. & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: Jonh Wiley & Sons, Inc.
- Flouri, E., Buchanan, A., & Bream, V. (2002). Adolescents' perceptions of their fathers' involvement: Significance to school attitudes. *Psychology in the Schools*, 39, 575-582
- Formiga, N. S. (2004). O tipo de orientação cultural e sua influência sobre os indicadores do rendimento escolar. *Psicologia: Teoria e Prática*, 6(1), 13-29.
- Goldstein, H. (1995) *Multilevel statistical models*. London: Edward Arnold.

- Goldstein, H. (1997). Value added tables: the less-than-holy grail. *Managing Schools Today*, 6, 18-19.
- Goldstein, H. (2001). The difficulty of ranking schools: The Limits to 'Value-added'. *New Economy*, 8(4), 197-198.
- Goldstein, H., Rasbasch, J., Yang, M., Geoffrey, W., Pan, H., Nuttall, D., & Thomas, S. (1993). A multilevel analysis of school examination results. *Oxford Review of Education*, 19(4), 425-433.
- Goldstein H., & Woodhouse G. (2008). Pesquisa sobre eficácia escolar e políticas educacionais. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 411-424). Belo Horizonte: Editora UFMG.
- Gomes, C. M. A. (2005). *Uma análise dos fatores cognitivos mensurados pelo Exame Nacional do Ensino Médio (ENEM)*. Tese de Doutorado, Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Gondim, S. M. G. (2002). Perfil profissional e mercado de trabalho: relação com formação acadêmica pela perspectiva de estudantes universitários. *Estudos de psicologia*, 7(2), 299-309. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-294X2002000200011&lng=en&nrm=iso>. Acesso em: 09 ago. 2009.
- Gray, J. (2008). Desenvolvendo métodos de valor agregado para avaliação da escola. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 252-260). Belo Horizonte: Editora UFMG.
- Gregory, R. (1994). Seeing intelligence. Em J. Khalfa (Ed). *What is intelligence?* (pp. 13-26). Cambridge, England: Cambridge University Press.

- 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.
- Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. Em K. S. McGrew, J. K. Werder, & R. W. Woodcock (Orgs.). *Woodcock Johnson technical manual* (pp. 197-232). Chicago: Riverside.
- Horn, J. L. & Cattell, R. B. (1966) Refinement and test of the theory of fluid and crystalized general intelligences. *Journal of Educational Psychology*, 57(5), 253-270.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107–129.
- Hunt, E. (2000). Let's hear it for crystallized intelligence. *Learning and Individual Differences*, 12, 123-129.
- Impara, J. C., & Plake, B. S. (2000). *A Comparison of Cut Scores using Multiple Standard Setting Methods*. Paper presented at the Large Scale Assessment Conference, Snowbird, UT.
- Jencks, C. (2008). Desigualdade no aproveitamento educacional. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 50-66). Belo Horizonte: Editora UFMG.
- Jesus, G. R. de. (2004). *Fatores que afetam o desempenho em português: um estudo multinível com dados do SAEB 2001*. Dissertação de Mestrado, Universidade de Brasília, Brasília.

- Jesus, G. R. & Laros, J. A. (2004). Eficácia escolar: regressão multinível com dados de avaliação em larga escala. *Avaliação Psicológica*, 3(2), 93-106.
- Kane, M. (1994) Validating the performance standards associated with passing scores. *Review of Educational Research*, 64 (3), p. 425-461.
- Kaufman, D. M., Mann, K. V., Muijtjens, A. M. M. & Vleuten, C. P. M. (2000). A Comparison of Standard-setting Procedures for an OSCE in Undergraduate Medical Education. *Academic Medicine*, 75(3), 267-271.
- Krutetskii, V. A. (1976). *The Psychology of Mathematical Abilities in Schoolchildren*. Chicago: The University of Chicago Press.
- Landeira-Fernandez, J., & Primi, R. (2002). Comparação do desempenho entre calouros e formandos no Provão de Psicologia 2000. *Psicologia: Reflexão e Crítica*, 15(1), 219-234.
- Laros, J. A. (2002). *Unidimensionalidade das provas do SAEB 1999*. Em VIII Conferência Internacional de Avaliação Psicológica - Formas e Contextos, V Encontro Mineiro de Avaliação Psicológica - Teorização e Prática - Anais. São Paulo: Vetor, 1. 166-168.
- Lee, V. E. (2008). Utilização e modelos hierárquicos lineares para estudar contextos sociais. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 273-296). Belo Horizonte: Editora UFMG.
- Leite, D. B. C. (2010). Estudantes e avaliação. *Avaliação*, 15(3), 9-27.
- Limana, A., & Brito, M. R. F. (2005). O modelo de avaliação dinâmica e o desenvolvimento de competências: algumas considerações a respeito do ENADE. *Revista Avaliação*, 10(2), 9-32.

- Linacre, J. M., & Wright, B. D. (1991). *WINSTEPS - Rasch-Model computer programs*. Chicago: MESA Press.
- Lopes, F. L., Vendramini, C. M. M., & Silva, M. C. R. (2010). O perfil do estudante de Pedagogia no Brasil. *Revista da Sociedade de Psicologia do Triângulo Mineiro*, 14(1).
- MacCann, R. G., & Stanley, G. (2006). The use of Rasch Modeling to improve Standard Setting. *Practical Assessment Research and Evaluation*, 11(2). Disponível em: <<http://pareonline.net/>>. Acesso em: 10 mai. 2008.
- Machado, M. D. (2003). A argumentação dos alunos egressos do ensino médio - um estudo das redações da prova do ENEM. *Dissertação de Mestrado*. Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara.
- Madaus, G. F., Airasian, P. W., & Kellaghan, T. (2008). Insumos escolares, processos e recursos. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 112-141). Belo Horizonte: Editora UFMG.
- Maggio, I. P. (2006). As políticas públicas de avaliação: o ENEM, expectativas e ações dos professores. *Dissertação de mestrado*. Pontifícia Universidade Católica, São Paulo.
- Mayer, J. & Salovey, P. (1998) O que é inteligência emocional? Em P. Salovey & D.J. Sluyter (Orgs.), *Inteligência emocional na criança: aplicações na educação e no dia-a-dia* (pp. 13-49). Rio de Janeiro: Campus.
- McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive *Gf-Gc* framework. Em D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151-179). New York: Guilford.

- McGrew, K. S. (2003). Cattell-Horn-Carroll (CHC). *Definition Project: Institute for Applied Psychometrics*. [on line] Disponível em: <<http://www.iapsych.com/chcdef.htm>>. Acesso em: 2 jun 2008.
- McGrew, K.S. (2004) Cattell-Horn-Carroll (CHC) Definition Project [on line]. Disponível em: <<http://www.iapsych.com/HCARR2.pdf>>. Acesso em: 2 jun 2008.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll (CHC) theory of cognitive abilities: Past, present and future. Em D. Flanagan, & Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (p.136-202). New York: Guilford Press.
- McGrew, K. S. & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment*. Needham Heights: Allyn & Bacon.
- Messick, S. (1995). Validity of psychological assessment: validation inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Ministério da Educação e Cultura (2004a) – Instituto Nacional de Estudos e pesquisas Educacionais Anísio Teixeira (INEP). *LEI Nº 10.861, de 14 de Abril de 2004*. Disponível em: <www.inep.gov.br/superior/enade>. Acesso em: 23 abr 2008.
- Ministério da Educação e Cultura (2004b) – Instituto Nacional de Estudos e pesquisas Educacionais Anísio Teixeira (INEP). *PORTARIA Nº 2.051, de 09 de Julho de 2004*. Disponível em: <www.inep.gov.br/superior/enade>. Acesso em: 23 abr 2008.
- Ministério da Educação e Cultura (2005) – Instituto Nacional de Estudos e pesquisas Educacionais Anísio Teixeira (INEP). *Matriz de Avaliação*. Disponível em: <<http://www.inep.gov.br/download/enade/2005/relatorios/>>. Acesso em: 23 abr 2008.

- Mildner, T., & Silva, A. (2002). O ENEM como forma alternativa ou complementar aos concursos vestibulares no caso das áreas de conhecimento “Língua Portuguesa e Literatura”: relevante ou passível de refutação? *Estudos em Avaliação Educacional*, 25, 43-75.
- Miranda, S. (2006). How the information needs may be related to information competency. *Ciencia da Informação*, 35(3), 99-114.
- Moreno, M. L. R. (2006). De la Evaluación a la Formación de Competencias Genéricas: Aproximación a un Modelo. *Revista Brasileira de Orientação Profissional*, 7(2), 33- 48.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (2008a). Por que fazer um estudo das escolas primárias? Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 153-162). Belo Horizonte: Editora UFMG.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (2008b). A necessidade de se considerar as características do estudante. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 187-215). Belo Horizonte: Editora UFMG.
- Muñoz-Repiso, M. & Murillo, F. J. (Coords.). (2003). *Mejorar las escuelas, mejorar los resultados. Investigación europea sobre mejora de la eficacia escolar*. Bilbao: Mensajero.
- Nogueira, S. O. (2008). *ENADE: Análise de Itens de Formação Geral e de Estatística pela TRI*. Dissertação de Mestrado, Universidade São Francisco, Itatiba, São Paulo.
- Nunes, C. H. S. S. & Primi, R (2010). Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos. Em Conselho Federal de Psicologia (CFP). (Org.). *Avaliação Psicológica: Diretrizes na regulamentação da profissão* (pp. 101-128). Brasília: CFP.

- Oliveira, Z. A. (2006). *Saberes e Práticas Avaliativas no Ensino de História: o impacto dos processos seletivos (PAIES e VESTIBULAR/UFU) e do ENEM na avaliação da aprendizagem no Ensino Médio*. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia.
- Oliveira, K. S. (2006). *Avaliação do Exame Nacional de Desempenho do Estudante pela Teoria de Resposta ao Item*. Dissertação de Mestrado, Universidade São Francisco, Itatiba, São Paulo.
- Oliveira, L. K. M., Franco, C., & Soares, T. M. (2007). Projeto GERES/ 2005: Novos indicadores para construção e interpretação da escala de proficiência. *Revista Eletrônica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)*, 5(2), 154-182.
- Opendakker, M. C. & Van Damme, J. (2000). *Effects of schools, teaching staff and classes on achievement and well-being in secondary education: similarities and differences between school outcomes*. *School Effectiveness and School Improvement*, 11, 165–196.
- Opendakker, M.C. and Van Damme, J. (2005). Enhancing effort and achievement: The importance of parent involvement and home-school partnerships. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 3(1). Disponível em: <http://www.ice.deusto.es/rinace/reice/vol3n1_e/OpendakkerVanDamme.pdf>. Acesso em: 23 abr. 2009.
- Paiva, G. S. (2008). Avaliação do desempenho dos estudantes da educação superior: a questão da equidade e obrigatoriedade no Provão e Enade. *Ensaio: aval. pol. públ. Educ.*, 16(58), 31-46.
- Parks, C.W. Jr., Mitchell, D.B., & Perlmutter, M. (1986). Cognitive and social functioning across adulthood: age or student status differences? *Psychol Aging*, 1(3), 248–254.

- Parry, S. B. (1996). The quest for competencies. *Training Magazine*, 33(7), 48-56.
- Parry, S. B. (1998). Just what is a competency? (And why should you care?). *Training*, 35(6), 58-64.
- Pasquali, L. (2001). *Psicometria: teoria dos testes psicológicos*. Brasília: LabPAM.
- Phelps, L., McGrew, K., Knopik, S. N., & Ford, L. (2005). The general (g), broad and narrow CHC stratum characteristics of the WJ III and WISC-III Tests: A confirmatory cross-battery investigation. *School Psychology Quarterly*, 20, 66-88.
- Primi, R. (1998). *Desenvolvimento de um instrumento informatizado para a avaliação do raciocínio analítico*. Tese Doutorado, Universidade de São Paulo, São Paulo.
- Primi, R. (2002a). Avanços na concepção psicométrica da inteligência. Em F. C. Capovilla (Org.). *Neuropsicologia e aprendizagem: uma abordagem multidisciplinar* (pp. 77-86). São Paulo: Scortecci, Sociedade Brasileira de Neuropsicologia.
- Primi, R. (2002b). Complexity of geometric inductive reasoning tasks. Contribution to the understanding of fluid intelligence. *Intelligence*, 30, 41-70.
- Primi, R. (2002c). Inteligência Fluida: Definição fatorial, cognitiva e neuropsicológica. *Cadernos de Psicologia e Educação - Paidéia*, 12 (2/3), 57-77.
- Primi, R. (2004). Avanços na Interpretação de Escalas com a Aplicação da Teoria de Resposta ao Item. *Avaliação Psicológica*, 3(1), 53-58.
- Primi, R. (2006a). Evidências de Validade das Provas do ENADE - 2004. Em D. Ristoff, A. Limana M. R. F. Brito (Orgs.). *ENADE: perspectiva de avaliação dinâmica e análise de mudanças* (pp. 59-76). Brasília: INEP/DAES/MEC.

- Primi, R. (2006b). *A Validade do ENADE para Avaliação da Qualidade dos Cursos de Instituições de Ensino Superior*. Projeto de Pesquisa. Itatiba: Universidade São Francisco, LabAPE.
- Primi, R., Santos, A. A. A., Vendramini, C. M. M., Taxa, F., Muller, A. M., Lukjanenko, M. F., & Sampaio, I. S. (2001). Competências e Habilidades Cognitivas: diferentes definições dos mesmos construtos. *Psicologia: Teoria e Pesquisa*, 17(2), 01-09.
- Primi, R., & Almeida, L. S. (2002). Inteligência geral ou fluida: Desenvolvimentos recentes na sua concepção. *Sobredotação*, 3, 127-144.
- Primi, R., Santos, A. A. A., & Vendramini, C. M. M. (2002). Habilidades básicas e desempenho acadêmico em ingressantes universitários. *Estudos de Psicologia*, 7(1), 47-55.
- Primi, R., Landeira-Fernandez, J., & Ziviani, C. (2003). O provão de psicologia: objetivos, problemas, conseqüências e sugestões. *Psicologia: Teoria e Pesquisa*, 19(2), 109-116.
- Primi, R. & Miguel, F.K. (2006). Análise da Prova de Medicina do ENADE pela Teoria de Resposta ao Item. Relatório Técnico não publicado.
- Primi, R., Nunes, C. H. S. S., Vendramini, C. M. M., Silva, M. C. R., Miguel, F. K., Dias, A. S., Carvalho, L. F., Nogueira, S. O. (2007a). *Análise Psicométrica da prova Pluridisciplinar do curso de Medicina e estudo comparativo com os resultados do ENADE 2004*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- Primi, R., Nunes, C. H. S. S., Vendramini, C. M. M., Silva, M. C. R., Miguel, F. K., Dias, A. S., Carvalho, L. F., Nogueira, S. O. (2007b). *Análise Psicométrica da prova Pluridisciplinar do curso de Educação Física e estudo comparativo com os resultados*

- do ENADE 2004. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- Primi, R., Couto, G., Almeida, L. S., & Guisande, M. A. (2008). Impacto da escolarização e da idade no desempenho em testes de raciocínio: Diferenças em função do conteúdo dos itens à luz da teoria gf-gc (Submetido). *Spanish Journal of Psychology*.
- Primi, R., Vendramini, C.M.M., Nunes, C.H.S.S., Silva, M. C. R., Bartholomeu, D., Pires, S. D., Mata, A.S., Carvalho, L.F., & Miguel, F.K. (2008a). *Análise Psicométrica e Resultados dos Alunos da USF Direito*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- 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., & Mata, A.S. (2008b). *Análise Psicométrica e Resultados do Curso de Administração da USF*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba
- Primi, R., Vendramini, C.M.M., Nunes, C.H.S.S., Carvalho, L.F., Miguel F.K., Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Mata, A.S. (2008c). *Análise Psicométrica e Resultados dos Alunos da USF Psicologia*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba
- Primi, R; Muniz, M, & Nunes, C. H. S. S. (2009). Definições Contemporâneas de Validade de Testes Psicológicos. Em C. S. Hutz (Org.). *Avanços e polêmicas em avaliação psicológica*. (pp. 243-265). São Paulo: Casa do Psicólogo

- Primi, R., Nunes, C. H. S. S., Silva, M. C. R., Carvalho, L. F., Miguel, F. K. & Vendramini, C. M. M. (2009a). Aplicação da Teoria de Resposta ao Item na Interpretação das Notas do ENADE de Psicologia. *Revista de Educação AEC*, 38(150), 115-124.
- Primi, R., Silva, M. C. R., Bartholomeu, D., Vendramini, C. M. M., Nunes, C. H. S. S. & Mata, A. S. (2009b). Questões metodológicas referentes ao Exame Nacional de Desempenho dos Estudantes (ENADE). *Revista de Educação AEC*, 38(150), 125-134.
- Primi, R. & Nunes, C. H. S. S. (2010). O SATEPSI: propostas de aprimoramento. Em Conselho Federal de Psicologia (CFP). (Org.). *Avaliação Psicológica: Diretrizes na regulamentação da profissão* (p. 129-148). Brasília: CFP.
- Primi, R., Carvalho, L. F., Miguel, F. K., & Silva, M. C. R. (no prelo). Análise do Funcionamento Diferencial dos Itens do Exame Nacional do Estudante (ENADE) de Psicologia de 2006. *Psico-USF*.
- Primi, R., Hutz, C. S., & Silva, M. C. R. (no prelo). *A prova do ENADE de Psicologia 2006: concepção, construção e análise da prova*.
- Primi, R., & Nunes, C. H. S. S. (no prelo). *Validação de provas educacionais em larga escala*.
- Raudenbush, S. W. (2004a). *Schooling, statistics, and poverty: can we measure school improvement?* New Jersey: Educational Testing Service.
- Raudenbush, S. W. (2004b). What are value-added models of estimating and what does this imply for statistical practice. *Journal of Educational and Behavioral Statistics*, 29(1), 121-129.
- Raudenbush, S. W., & Wilms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307-335.

- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Advanced Quantitative Techniques in the Social Sciences). London: Sage Publications.
- Rauter, R., Laros, J. A., & Jesus, G. R. (2007). Questionários contextuais do SAEB: análise qualitativa baseada nos resultados da análise fatorial. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(2), 56-68.
- Real, G. C. M. (2007). *A qualidade revelada na educação superior: impactos da política de avaliação no Brasil*. Tese de Doutorado, Faculdade de Educação, Universidade de São Paulo, São Paulo.
- Riani, J. L. R. (2005). *Determinantes do resultado educacional no Brasil: família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial*. Tese de Doutorado, Belo Horizonte, UFMG/CEDEPLAR
- Ribeiro, I. (1998). *Mudanças no desempenho e na estrutura das aptiões: Contributos para o estudo da diferenciação cognitiva em jovens*. Braga: Universidade do Minho.
- Rodrigues, J. (1994) *A taxonomia de objetivos educacionais - um manual para o usuário*. Editora UNB.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701.
- Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, 29, 103-116.
- Rutter, M. (1979) *Changing youth in a changing society: patterns of adolescent development and disorder*. London: Nulfield Provincial Hospitals Trust.

- Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (2008a). Estudos anteriores. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 142-152). Belo Horizonte: Editora UFMG.
- Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (2008b). A busca pela eficácia. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 153-186). Belo Horizonte: Editora UFMG.
- Salthouse T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, 2, 179-183.
- Sanders, W. L. & Horn, S P. (1995). The Tennessee Value-Added Assessment System (TVAAS): Mixed Model Methodology in Educational Assessment. Em Shrinkfield, A. J. & Stufflebeam, D. (Eds.), *Teacher Evaluation: Guide to Effective Practice* (pp. 337-350). Boston, MA: Kluwer.
- Santos Neto, A. D. (2006). *O processo de contextualização nas escolas públicas de ensino médio do DF com desempenho acima da média no Exame Nacional do Ensino Médio*. Dissertação de Mestrado, Universidade Católica de Brasília.
- Schelini, P. W., & Wechsler, S. (2006). Estudo da estrutura fatorial da bateria multidimensional de inteligência infantil. *Estudos de Psicologia*, 23(2), 105-112.
- Schölkopf, D., Smola A.J., & Müller, K. R. (1999). Kernel Principal Component Analysis. Em B. Scholkopf, C.J.C. Burgers & A. J. Smola (Orgs.), *Advances in Kernel Methods--Support Vector Learning* (pp. 293-306). Cambridge, MA: MIT Press.
- Schooler, C. (1987). Cognitive effects of complex environments during the life span: A review and theory. Em C. Schooler & K.W. Schaie (Eds.). *Cognitive functioning and social structure over the life course* (pp. 24-49). Norwood, NJ: Ablex.

- Schwartzman, S. (2010). *O Impacto dos rankings nas instituições de ensino*. Trabalho apresentado no VIII Congresso Brasileiro de Gestão Educacional, São Paulo. Disponível em: <<http://www.schwartzman.org.br/simon/2010rankings.pdf>>. Acessado em: 10 ago. 2010.
- Schweizer, K., & Koch, W. (2002). A revision of Cattell's Investment Theory: Cognitive properties influencing learning. *Learning and Individual Differences*, 13, 57–82.
- Searle, J. (2000). Defining competency: the role of standard setting. *Medical Education*, 34, 363-366.
- Shen, L. (2001). *A comparison of Angoff and Rasch Model based Item Map Methods in Standard Setting*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA. Disponível em: <http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED452213&ERICExtSearch_SearchType_0=no&accno=ED452213>. Acesso em: 15 mai 2009.
- Silva, M. C. R., Dias, A. S., Vendramini, C. M. M., & Brito, M. R. F. (2007). *Contribuição da Iniciação Científica para o desempenho dos estudantes no ENADE*. III Congresso Brasileiro de Avaliação Psicológica e XII Conferência Internacional da Avaliação Psicológica: Formas e Contextos, João Pessoa.
- Silva, M. C. R., Dias, A. S., Carvalho, L. F., Primi, R., Nunes, C. H. S. S., Vendramini, C. M. M., Miguel, F. K., Nogueira, S. O. (2007). Análise Psicométrica da prova Pluridisciplinar do curso de Farmácia e estudo comparativo com os resultados do ENADE 2004. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.

- Silva, M. C. R., Bartholomeu, D., Primi, R., Vendramini, C. M. M., Pires, S. D., Carvalho, L. F., & Nunes, C. H. S. S. (2010). Proposta de organização das variáveis do questionário socioeconômico do ENADE para pesquisas em eficácia escolar. *Revista da Sociedade de Psicologia do Triângulo Mineiro*, 14(2).
- Silva, M. C. R., Bartholomeu, D., Pires, S. D., & Primi, R. (2010). Teoria de Resposta ao Item: equalização da prova de Administração do ENADE. Em F. C. Capovilla. (Org.). *Transtornos de aprendizagem: Progressos em avaliação e intervenção preventiva e remediativa* (p. 239-246), v. 1, São Paulo: Memnon.
- Silva, M. C. R., Vendramini, C. M. M. & Lopes, F. L. (2010). Diferenças entre gênero e perfil sócio-econômico no Exame Nacional de Desempenho do Estudante. *Avaliação*, 15(3), 185-202.
- Smith, J., & Marsiske, M. (1994). *Abilities and competencies in adulthood: Life-Span perspectives on workplace skills*. Paris: Organization for Economic Co-Operation and Development; Centre for Educational Research and Development.
- Smith Jr., E. V. (2004). Metric Development and Score Reporting in Rasch Measurement. Em E. V. Smith Jr. & R. M. Smith, (Orgs.), *Introduction to Rasch Measurement: Theory, Models, and Applications* (pp.). JAM Press, Maple Glover.
- Smith Jr., E. V., & Smith, R. M. (eds.) (2004). *Introduction to Rasch Measurement: Theory, Models, and Applications*. JAM Press, Maple Glover.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling*. London: Sage Publications.
- Snijders, T. A. B., & Bosker, R. (2003). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London: Sage Publications.

- Snow, R.E. (1992). Aptitude theory: Yesterday, today and tomorrow. *Educational Psychologist*, 27, 5-32.
- Soares, J. F. (2004). O efeito da escola no desempenho cognitivo de seus alunos. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)*, 2(2). Disponível em: <<http://www.ice.deusto.es/rinace/reice/vol2n2/Soares.pdf>>. Acesso em: 20 mai. 2008. Acesso em: 15 mai 2009.
- Soares, J. F. (2005). O efeito da escola no desempenho cognitivo de seus estudantes. *Revista Electrónica Iberoamericana sobre Calidad, Eficácia y Cambio em Educación*, 2(2), 83-104.
- Soares, J. F. (2007). Melhoria do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de Pesquisa*, 37(130), 135-160.
- Soares, J. F. (2005). O efeito da escola no desempenho cognitivo de seus alunos. Em A. M. Souza (Org.), *Dimensões da avaliação educacional* (pp.174-204). Petrópolis: Vozes.
- Soares, J. F., Castro, C. M., & Ribeiro, L. M. (2000). *O Provão: os cursos A são os que mais oferecem aos estudantes?* Belo Horizonte: UFMG. Texto não publicado.
- Soares, J. F., Sátyro, N. G. D. & Mambrini, J. (2000). *Modelo explicativo do desempenho escolar dos alunos e análise dos fatores do SAEB - 1997*. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas.
- Soares, J. F., Ribeiro, L. & 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.

- Soares, J. F., César, C. C., & Mambrini, J. (2001). Determinantes do desempenho dos alunos do ensino básico brasileiro: evidências do Saeb de 1997. Em C. Franco (Org.), *Promoção, ciclos e avaliação educacional* (pp. 121-153). Porto Alegre: Artmed.
- Soares, J. F., Mambrini, J., Pereira, L. R., & Alves, M. T. G. (2004). *Fatores associados ao desempenho em língua portuguesa e matemática: a evidência do SAEB-2003*. Belo Horizonte: Laboratório de Medidas Educacionais, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Relatório Técnico.
- Soares, J. F. & Andrade, R. J. (2006). Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. *Ensaio: aval. pol. públ. Educ.*, 14(50), 107-126.
- Soares, J. F. & Collares, A. C. M. (2006). Recursos Familiares e o Desempenho Cognitivo dos Alunos do Ensino Básico Brasileiro. *DADOS: Revista de Ciências Sociais*, 49(3), 615-650.
- Sousa, S. M. Z. L. (2003). Possíveis impactos das políticas de avaliação no currículo escolar. *Cadernos de Pesquisa*, 119, 175-190.
- 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>. Acesso em: 16 mar. 2009.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
- Sternberg, R. J. (1997). *Thinking styles*. New York, NY: Cambridge University Press.
- Sternberg, R. J. (2000). *Handbook of Intelligence*. New York: Cambridge University Press.

- Sternberg, R. J., & Grigorenko, E. L. (2003). *The psychology of abilities, competencies, and expertise*. New York: Cambridge University Press.
- Stevão, C. B. G. A. (2008). *Teoria da Resposta ao Item: Um estudo inicial os dados GERES Campinas*. Dissertação de Mestrado, Faculdade de Educação, Universidade Estadual de Campinas, São Paulo.
- Torrecilla, F. J. M. (2008). Um panorama da pesquisa ibero-americana sobre a eficácia escolar. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 466-481). Belo Horizonte: Editora UFMG.
- Ubriaco, F. E. C. A. (2009) *Interpretação de Escalas de Medida de Competência Matemática*. Dissertação de Mestrado, Faculdade de Educação, UFMG.
- Undheim, J. O. (1981). On intelligence IV: toward a restoration of general intelligence. *Scandinavian Journal of Psychology*, 22, 251-266.
- Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed
- Vendramini, C. M. M. (2002). Aplicação da teoria de resposta ao item na avaliação educacional. Em Primi R. (Org.). *Temas em avaliação psicológica* (pp. 116-130). Campinas: Instituto Brasileiro de Avaliação Psicológica.
- Vendramini, C. M. M. (2005). O Uso de um Modelo Multidimensional para Análise do ENADE. *Revista de Avaliação Institucional*, 10(3), 27-40.
- Vendramini, C. M. M. (2007). *Propriedades psicométricas das provas do ENADE 2004 e 2005*. Trabalho apresentado no III Congresso Brasileiro de Avaliação Psicológica e XII Conferência Internacional de Avaliação Psicológica: Formas e Contextos, João Pessoa, PB.

- Vendramini, C. M. M., Nunes, C. H. S., Primi, R., S., Nogueira, S. O., Silva, M. C. R., Lopes, F. L., Dias, A. S., Miguel, F. K., Carvalho, L. F. (2007a). *Análise Psicométrica da prova Pluridisciplinar do curso de Odontologia e estudo comparativo com os resultados do ENADE 2004*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- Vendramini, C. M. M., Nunes, C. H. S., Primi, R., S., Nogueira, S. O., Silva, M. C. R., Lopes, F. L., Dias, A. S., Miguel, F. K., Carvalho, L. F. (2007b). *Análise Psicométrica da prova Pluridisciplinar do curso de Fisioterapia e estudo comparativo com os resultados do ENADE 2004*. Relatório Técnico não publicado, Laboratório de Avaliação Psicológica e Educacional, Universidade São Francisco, Itatiba.
- Vendramini, C. M. M. (2009). Análise Fatorial baseada na TRI do componente de formação geral do ENADE. *Revista de Educação ANEC*, 38(150), 153-163.
- Vendramini, C. M. M., Silva, M. C. R., Dias, A. S., & Nogueira, S. O. (2007). *Evidências de validade da prova de pedagogia do ENADE 2005*. Pôster apresentado no III Congresso Brasileiro de Avaliação Psicológica e XII Conferência Internacional de Avaliação Psicológica: Formas e Contextos, João Pessoa, PB.
- 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: 30 ago 2008.
- Verhine, R. E., Dantas, L. M. V., & Soares, J. F. (2006). Do Provão ao Enade: uma análise comparativa dos exames nacionais utilizados no Ensino Superior brasileiro. *Ensaio: Avaliação e Políticas Públicas em Educação*, 14(52), 291-310.

- Vianna, H. M. (2003). Avaliações nacionais em larga escala: análises e propostas. *Estudos em Avaliação Educacional*, 27, 41-76.
- Voelkle, M. C., Wittmann, W. W., & Ackerman, P. L. (2006). Abilities and skill acquisition: A latent growth curve approach. *Learning and Individual Differences*, 16, 303–319
- Willms, J. D. (2008). A estimação do efeito da escola. Em N. Brooke & J. F. Soares, (Org.). *Pesquisa em eficácia escolar: origem e trajetórias* (pp. 261-272). Belo Horizonte: Editora UFMG.
- Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment*, 8, 231-258.
- Wright, B. D., & Stone, M. H. (1979). Best test design. Chicago: MESA Press.
- Zanchet, B. M. B. A. (2003). *A prática avaliativa do exame nacional do ensino médio (ENEM): pressupostos conceituais e implicações no contexto escolar*. Tese de Doutorado, Universidade do Vale do Rio dos Sinos.
- Zirondi, M. I. (2007). *Desvendando aspectos de linguagem no ENEM: uma contribuição para o processo educacional brasileiro*. Dissertação de Mestrado, Universidade Estadual de Londrina, Londrina.
- Ziviani, C. & Primi, R. (2002). Teoria de Resposta ao Item e o modelo de Rasch de mensuração: uma análise do Provão de Psicologia. Em R. Primi (org.). *Temas em Avaliação Psicológica* (pp. 131-151). Campinas: IDB Digital / Instituto Brasileiro de Avaliação Psicológica (IBAP).

11. ANEXOS

Anexo 1 – Aprovação do Comitê de Ética



Comitê de Ética em Pesquisa – CEP

Bragança Paulista, 22 de Outubro de 2009.

COMITÊ DE ÉTICA EM PESQUISA

Projeto de Pesquisa: Validade do ENADE por meio do Modelo de Rasch: avaliação de competências.

Curso: Psicologia

Grupo: III

Autor (es): Prof. Dr. Ricardo Primi; Marjorie Cristina Rocha da Silva.

Instituição: UNIVERSIDADE SÃO FRANCISCO

Protocolo : 0233.0.142.000-09 (Citar este número nas correspondências referentes a este projeto)

Prezado (a) (s) Pesquisador (a) (s),

O Comitê de Ética em Pesquisa – CEP, da Universidade São Francisco, analisou em reunião extraordinária no dia **22/10/2009** o projeto de pesquisa supracitado, sob a responsabilidade de Vossa Senhoria.

Este Comitê, acatando o parecer do relator indicado, apresenta-lhe o seguinte resultado:

Parecer: APROVADO

- **Sugestão:** O pedido ao Comitê deve se referir à Legislação do SISNEP e não do Conselho Federal de Psicologia.

Atenciosamente,

Patrícia de Oliveira Carvalho
Vice- Coordenadora do Comitê de Ética em Pesquisa - USF
Universidade São Francisco

Ao término do desenvolvimento do estudo o (a) pesquisador (a) responsável deverá enviar ao CEP-USF o relatório consolidado de conclusão do mesmo.

Anexo 2 - Objetivos, diretrizes e conteúdos propostos na avaliação do curso de Administração

O Exame Nacional de Desempenho dos Estudantes (ENADE/2006), no componente específico da área de Administração, teve por *objetivos*:

- a) acompanhar o processo de aprendizagem e o desempenho dos estudantes de Administração;
- b) verificar a aquisição dos conhecimentos, habilidades e competências necessários ao pleno exercício da profissão e da cidadania;
- c) contribuir para o processo de avaliação dos cursos e das instituições formadoras na área de Administração;
- d) subsidiar o processo de avaliação institucional dos cursos de graduação em Administração;
- e) consolidar o processo de avaliação do desempenho dos estudantes como um componente do Sistema Nacional de Avaliação da Educação Superior (SINAES).

A prova do ENADE/2006, no componente específico da área de Administração, avaliou se o estudante desenvolveu, durante sua formação, as seguintes *habilidades e competências*:

- a) reconhecer e definir problemas, equacionar soluções, pensar estrategicamente, introduzir modificações no processo produtivo, atuar preventivamente, transferir e generalizar conhecimentos e exercer, em diferentes graus de complexidade, o processo da tomada de decisão;
- b) desenvolver expressão e comunicação compatíveis com o exercício profissional, inclusive nos processos de negociação e nas comunicações interpessoais ou intergrupais;

- c) refletir e atuar criticamente sobre a esfera da produção, compreendendo sua posição e função na estrutura produtiva sob seu controle e gerenciamento;
- d) desenvolver raciocínio lógico, crítico e analítico para operar com valores e formulações matemáticas presentes nas relações formais e causais entre fenômenos produtivos, administrativos e de controle, bem assim expressando-se de modo crítico e criativo diante dos diferentes contextos organizacionais e sociais;
- e) ter iniciativa, criatividade, determinação, vontade política e administrativa, vontade de aprender, abertura às mudanças e consciência da qualidade e das implicações éticas do seu exercício profissional;
- f) desenvolver capacidade de transferir conhecimentos da vida e da experiência cotidianas para o ambiente de trabalho e do seu campo de atuação profissional, em diferentes modelos organizacionais, revelando-se profissional adaptável;
- g) desenvolver capacidade para elaborar, implementar e consolidar projetos em organizações; e
- h) desenvolver capacidade para realizar consultoria em gestão e administração, pareceres e perícias administrativas, gerenciais, organizacionais, estratégicos e operacionais.

A prova do ENADE/2006, no componente específico da área de Administração, adotou como referencial os seguintes conteúdos:

- a) conteúdos de Formação Básica: relacionados com estudos antropológicos, sociológicos, filosóficos, psicológicos, ético-profissionais, políticos, comportamentais, econômicos e contábeis, bem como os relacionados com as tecnologias da comunicação e da informação e das ciências jurídicas;
- b) conteúdos de Formação Profissional: relacionados com as áreas específicas, envolvendo teorias da administração e das organizações e a administração de recursos humanos,

mercado e *marketing*, materiais, produção e logística, financeira e orçamentária, sistemas de informações, planejamento estratégico e serviços;

c) conteúdos de Estudos Quantitativos e suas Tecnologias: abrangendo pesquisa operacional, teoria dos jogos, modelos matemáticos e estatísticos e aplicação de tecnologias que contribuam para a definição e utilização de estratégias e procedimentos inerentes à Administração.

Tabela 20 - *Habilidades Aferidas nas questões objetivas da FG do ENADE 2006 de Administração*

Item	Conteúdos predominantes	Habilidades aferidas
1	Políticas públicas: educação	▪ Estabelecer comparações
2	Arte e Literatura	▪ Ler e compreender texto, estabelecer relações entre gravura e texto
3	Multiculturalismo	▪ Analisar argumento
4	Ética	▪ Analisar situações-problema
5	Sociodiversidade	▪ Fazer inferências por dedução
6	Políticas públicas: saúde e segurança	▪ Compreender informações, estabelecer relações contexto-situação
7	Relações de trabalho	▪ Compreender informações, analisar tabela
8	Arte e Filosofia	▪ Ler e compreender texto verbal e não verbal, estabelecer relações

Tabela 21 - *Habilidades Aferidas nas questões objetivas do CE do ENADE 2006 de Administração*

Item	Conteúdos predominantes	Habilidades aferidas
11	Teorias da Administração	▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)
12	Teorias da Administração	▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)

Tabela 21 - Continuação

Item	Conteúdos predominantes	Habilidades aferidas
13	Teorias da Administração, Planejamento	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Compreensão do todo administrativo e de suas relações com o ambiente externo (4)
14	Teorias da Administração, Administração de Recursos Humanos e Psicologia.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)
15	Teorias da Administração, Psicologia e Sociologia	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação de problemas organizacionais (3) ▪ Interação criativa com os diferentes conflitos organizacionais (5)
16	Administração de Recursos Humanos, Psicologia e Sociologia.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Seleção de estratégias adequadas de ação, visando e atender interesses interpessoais (6)
17	Administração de Recursos Humanos	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da organização (7)
18	Administração de Recursos Humanos	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da organização (7)

Tabela 21 – Continuação

Item	Conteúdos predominantes	Habilidades aferidas
19	Teorias da Administração e Administração Mercadológica.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Interação criativa com os diferentes interesses organizacionais (5) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da organização (7)
20	Administração Mercadológica.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Interação criativa com os diferentes interesses organizacionais (5) ▪ Seleção de estratégias adequadas de ação, visando e atender interesses institucionais (6) ▪ Resolução de problemas organizacionais com flexibilidade e adaptabilidade (8)
21	Administração Mercadológica e Economia.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Seleção de estratégias adequadas de ação, visando e atender interesses interorganizacionais e institucionais (6) ▪ Resolução de problemas organizacionais com flexibilidade e adaptabilidade (8) ▪ Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões (9)
22	Administração Mercadológica	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na identificação de problemas organizacionais (3) ▪ Resolução de problemas organizacionais com flexibilidade e adaptabilidade (8) ▪ Seleção de estratégias adequadas de ação, visando e atender interesses interorganizacionais e institucionais (6)

Tabela 21 – Continuação

Item	Conteúdos predominantes	Habilidades aferidas
23	Administração Estratégica	<ul style="list-style-type: none"> ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na solução de problemas organizacionais (3) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos da organização (7) ▪ Identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Percepção de modelos de estratégia e seu impacto na gestão (10)
24	Administração Financeira e Orçamentária, Contabilidade e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2)
25	Administração Financeira e Orçamentária, Economia, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Hierarquização de atividades e programas para tomada de decisões (9)
26	Administração Financeira e Orçamentária, Administração de Produção e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Identificação e dimensionamento de riscos para tomada de decisões (9)
27	Administração de Serviços	<ul style="list-style-type: none"> ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico na solução de problemas organizacionais (3) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos da organização (7) ▪ Tomada de decisão em serviços (9) ▪ Percepção de modelos de gestão de serviços (10)
28	Administração de Sistemas de Informação e Tecnologia da Informação	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio crítico na solução de problemas organizacionais (3) ▪ Percepção de modelos inovadores de gestão (10)
29	Administração de Sistemas de Informação e Tecnologia da Informação	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio crítico na solução de problemas organizacionais (3) ▪ Percepção de modelos inovadores de gestão (10)

Tabela 21 – Continuação

Item	Conteúdos predominantes	Habilidades aferidas
30	Administração de Produção, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Resolução de problemas organizacionais com flexibilidade e adaptabilidade (8)
31	Administração de Produção, Organização e Processos, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Hierarquização de atividades e programas para tomada de decisões (9) ▪ Compreensão dos processos produtivos e controle estatístico dos processos (11)
32	Administração de Produção, Processos, Logística, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Hierarquização de atividades e programas para tomada de decisões (9) ▪ Percepção de modelos inovadores de gestão (10) ▪ Compreensão dos processos produtivos e controle estatístico dos processos (11) ▪ Entendimento dos canais de distribuição e cadeia de suprimento (12)
33	Administração de Recursos Materiais e Patrimoniais e Logística	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Hierarquização de atividades e programas para tomada de decisões (3) ▪ Entendimento dos canais de distribuição e cadeia de suprimento (12)
34	Administração de Recursos Materiais e Patrimoniais, Estatística e Matemática	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Compreensão dos processos produtivos e controle estatístico dos processos (11)

Tabela 22 - Habilidades Aferidas nas questões dissertativas do CE de Administração

Item	Conteúdos predominantes	Habilidades aferidas
35	<p>Administração Financeira e Orçamentária, Administração de Sistemas de Informação, Administração de Produção, Administração de Recursos Materiais e Patrimoniais, Ética e Responsabilidade Social, Contabilidade, Economia, Matemática, Estatística, Tecnologia da Informação e Logística. (Estudo de Caso)</p>	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (3) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da administração (7) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)
36	<p>Administração Financeira e Orçamentária, Administração de Sistemas de Informação, Administração de Produção, Administração de Recursos Materiais e Patrimoniais, Ética e Responsabilidade Social, Contabilidade, Economia, Matemática, Estatística, Tecnologia da Informação e Logística. (Estudo de Caso)</p>	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (3) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da administração (7) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)

Tabela 22 - Continuação

37	<p>Administração Financeira e Orçamentária, Administração de Sistemas de Informação, Administração de Produção, Administração de Recursos Materiais e Patrimoniais, Ética e Responsabilidade Social, Contabilidade, Economia, Matemática, Estatística, Tecnologia da Informação e Logística. (Estudo de Caso)</p>	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (3) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da administração (7) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)
38	<p>Administração Financeira e Orçamentária, Administração de Sistemas de Informação, Administração de Produção, Administração de Recursos Materiais e Patrimoniais, Ética e Responsabilidade Social, Contabilidade, Economia, Matemática, Estatística, Tecnologia da Informação e Logística. (Estudo de Caso)</p>	<ul style="list-style-type: none"> ▪ Operação com valores e formulações quantitativas, estabelecendo relações formais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (3) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da administração (7) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)

Tabela 22 - Continuação

39	Teorias da Administração, Administração Mercadológica, Psicologia, Sociologia, Antropologia, Filosofia e Economia.	<ul style="list-style-type: none"> ▪ Operação com formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (3) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)
40	Teorias da Administração, Administração de Recursos Humanos, Psicologia e Sociologia.	<ul style="list-style-type: none"> ▪ Raciocínio lógico e crítico na identificação e solução de problemas organizacionais (1) ▪ Articulação do conhecimento sistematizado com a ação profissional (2) ▪ Compreensão do todo administrativo de modo integrado, sistêmico e estratégico, bem como suas relações com o ambiente externo (4) ▪ Operação com formulações qualitativas, estabelecendo relações formais e causais entre fenômenos organizacionais. ▪ Interação criativa com os diferentes interesses organizacionais e sociais (5) ▪ Seleção de estratégias adequadas de ação, visando a atender interesses interorganizacionais e institucionais (6) ▪ Resolução de problemas e desafios organizacionais com flexibilidade e adaptabilidade (8) ▪ Identificação e dimensionamento de riscos para tomada de decisões (9) ▪ Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações (13) ▪ Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos (14)

Apesar das diretrizes apontarem para 8 habilidades e competências gerais, ao verificar as especificações de cada um dos itens no Componente Específico da prova

observa-se que na verdade são propostas as avaliações de 14 habilidades específicas, que foram assim agrupadas apesar de algumas diferenças nas denominações (ora mais gerais, ora mais detalhistas). Portanto, o ENADE se propõe também a medir um conjunto de habilidades e competências a partir da capacidade do aluno de resolver problemas dentro de uma ampla gama de conteúdos que estão presentes na matriz curricular de cada curso.

No curso de Administração, as questões se propõem a medir as seguintes Habilidades: 1- Operação com formulações qualitativas e quantitativas, estabelecendo relações formais entre fenômenos organizacionais; 2- Articulação do conhecimento sistematizado com a ação profissional; 3- Raciocínio lógico e crítico na identificação de problemas organizacionais; 4- Compreensão do todo administrativo e de suas relações com o ambiente externo; 5- Interação criativa com os diferentes conflitos organizacionais; 6- Seleção de estratégias adequadas de ação, visando e atender interesses interpessoais, interorganizacionais e institucionais; 7- Adoção de procedimentos administrativos que contribuam para o alcance dos objetivos comuns da organização; 8- Resolução de problemas organizacionais com flexibilidade e adaptabilidade; 9- Hierarquização de atividades e programas, identificação e dimensionamento de riscos para tomada de decisões; 10- Percepção de modelos de estratégia e seu impacto na gestão de serviços; 11- Compreensão dos processos produtivos e controle estatístico dos processos; 12- Entendimento dos canais de distribuição e cadeia de suprimento; 13- Utilização da linguagem com clareza, precisão, propriedade na comunicação, fluência verbal e riqueza de vocabulário nos documentos técnicos específicos, bem como nas relações interpessoais, de forma a auxiliar na interpretação da realidade das organizações; e 14- Apreensão, articulação e sistematização de conhecimentos teóricos e metodológicos.

Anexo 3 - Análise dos itens do Componente Específico

Tabela 23 - Categorias de resposta nos itens de Componente Específico

ITEM	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE MEASURE	S.E. MEAN	OUTFIT MNSQ	PTBSE CORR
CE_11	B	0	19462	15	-.68	.00	1.0	-.03
	A	0	33766	26	-.67	.00	1.0	-.05
	D	0	31022	24	-.67	.00	1.0	-.04
	C	0	14057	11	-.64	.00	1.0	.00
	E	1	31035	24	-.43	.00	1.0	.11
	MISSING ***		924	1#	-.62	.03		.01
CE_12	D	0	19002	15	-.82	.00	.9	-.12
	E	0	5743	4	-.77	.01	.9	-.04
	C	0	22261	17	-.75	.00	1.0	-.07
	A	0	28352	22	-.60	.00	1.1	.04
	B	1	53979	42	-.47	.00	1.0	.12
	MISSING ***		929	1#	-.65	.04		.00
CE_13	B	0	8566	7	-1.00	.01	.8	-.13
	E	0	4997	4	-1.00	.01	.8	-.10
	D	0	10434	8	-.91	.01	.9	-.11
	A	0	17295	13	-.80	.00	1.0	-.08
	C	1	87954	68	-.48	.00	.9	.24
	MISSING ***		1020	1#	-.75	.03		-.01
CE_14	B	0	1618	1	-1.28	.02	.7	-.09
	E	0	9551	7	-1.04	.01	.8	-.15
	A	0	5946	5	-.94	.01	.9	-.09
	C	0	9308	7	-.93	.01	.9	-.10
	D	1	102917	80	-.51	.00	.9	.24
	MISSING ***		926	1#	-.75	.03		.00
CE_15	C	0	3210	2	-1.13	.01	.7	-.11
	A	0	4148	3	-1.06	.01	.8	-.10
	D	0	11664	9	-.96	.01	.8	-.14
	B	0	21005	16	-.71	.00	1.1	-.02
	E	1	89313	69	-.50	.00	1.0	.18
	MISSING ***		926	1#	-.79	.04		-.01
CE_16	E	0	2224	2	-1.09	.01	.7	-.09
	D	0	4934	4	-.94	.01	.9	-.08
	C	0	6336	5	-.92	.01	.9	-.08
	B	0	28837	22	-.62	.00	1.2	.06
	A	1	87203	67	-.56	.00	1.0	.04
	MISSING ***		732	1#	-.62	.07		.02
CE_17	C	0	13489	10	-.82	.00	.9	-.11
	D	0	9078	7	-.69	.01	1.0	-.03
	E	0	39330	30	-.64	.00	1.0	-.04
	B	0	45862	35	-.61	.00	1.0	.02
	A	1	21648	17	-.39	.00	1.0	.13
	MISSING ***		859	1#	-.68	.05		.00

Tabela 23 - continuação

CE_18	D	0	4464	3	-.98	.01	.9	-.09
	C	0	7439	6	-.92	.01	.9	-.10
	E	0	15036	12	-.80	.00	1.0	-.08
	A	0	17665	14	-.77	.00	1.0	-.07
	B	1	84712	66	-.50	.00	1.0	.19
	MISSING ***		950	1#	-.76	.04		-.01
CE_19	E	0	7448	6	-.83	.01	.9	-.08
	C	0	21098	16	-.77	.00	1.0	-.09
	D	0	9517	7	-.74	.01	1.0	-.04
	B	0	28934	22	-.71	.00	1.0	-.06
	A	1	62411	48	-.46	.00	1.0	.17
	MISSING ***		858	1#	-.73	.05		.00
CE_20	B	0	4402	3	-.92	.01	.9	-.08
	D	0	29983	23	-.77	.00	.9	-.12
	A	0	26125	20	-.70	.00	1.0	-.05
	E	0	16033	12	-.64	.00	1.0	.00
	C	1	52831	41	-.44	.00	1.0	.18
	MISSING ***		892	1#	-.67	.05		.00
CE_21	B	0	12661	10	-.83	.01	.9	-.10
	C	0	33251	26	-.67	.00	1.0	-.04
	A	0	14414	11	-.60	.00	1.1	.02
	E	0	40353	31	-.57	.00	1.1	.07
	D	1	28695	22	-.51	.00	1.1	.02
	MISSING ***		892	1#	-.68	.05		.00
CE_22	E	0	7823	6	-.80	.01	.9	-.08
	C	0	11226	9	-.78	.01	.9	-.07
	D	0	28857	22	-.68	.00	1.0	-.03
	A	0	38972	30	-.66	.00	1.0	-.02
	B	1	42528	33	-.44	.00	1.0	.13
	MISSING ***		860	1#	-.75	.05		.00
CE_23	B	0	14591	11	-.73	.00	1.0	-.06
	D	0	42025	33	-.70	.00	.9	-.09
	A	0	20689	16	-.61	.00	1.1	.02
	E	0	25283	20	-.50	.00	1.2	.12
	C	1	26442	20	-.51*	.00	1.1	.01
	MISSING ***		1236	1#	-.62	.03		-.01
CE_24	A	0	22792	18	-.83	.00	.9	-.15
	B	0	22158	17	-.73	.00	1.0	-.08
	C	0	25768	20	-.72	.00	1.0	-.08
	D	0	12128	9	-.71	.00	1.0	-.04
	E	1	45622	36	-.36	.00	.9	.27
	MISSING ***		1798	1#	-.61	.02		-.02
CE_25	C	0	10638	8	-.94	.01	.8	-.16
	A	0	6396	5	-.94	.01	.8	-.12
	B	0	9585	7	-.88	.01	.8	-.11
	E	0	57380	45	-.60	.00	1.1	.05
	D	1	44862	35	-.44	.00	1.0	.15
	MISSING ***		1405	1#	-.72	.02		-.03

Tabela 23 - continuação

CE_26	C	0	26606	21	-.75	.00	.9	-.11
	D	0	19345	15	-.74	.00	.9	-.08
	A	0	24177	19	-.71	.00	1.0	-.07
	E	0	21864	17	-.53	.00	1.1	.09
	B	1	36559	28	-.42	.00	1.0	.14
	MISSING ***		1715	1#	-.58	.02		-.02
CE_27	D	0	4265	3	-.96	.01	.8	-.09
	E	0	4827	4	-.80	.01	.9	-.05
	B	0	27747	21	-.70	.00	1.0	-.05
	C	0	52992	41	-.57	.00	1.1	.08
	A	1	39408	30	-.54	.00	1.1	.02
	MISSING ***		1027	1#	-.72	.04		-.01
CE_28	D	0	10826	8	-.88	.01	.9	-.12
	B	0	15447	12	-.77	.00	1.0	-.08
	C	0	15414	12	-.77	.00	1.0	-.09
	A	0	26261	20	-.68	.00	1.0	-.03
	E	1	60924	47	-.45	.00	1.0	.20
	MISSING ***		1394	1#	-.63	.03		-.02
CE_29	C	0	13544	10	-.87	.00	.8	-.14
	B	0	22991	18	-.74	.00	1.0	-.09
	D	0	36305	28	-.66	.00	1.0	-.03
	E	0	12056	9	-.66	.01	1.0	-.01
	A	1	44240	34	-.41	.00	1.0	.19
	MISSING ***		1130	1#	-.71	.03		-.02
CE_30	A	0	11710	9	-.80	.00	.9	-.09
	E	0	37861	29	-.78	.00	.9	-.16
	D	0	11021	9	-.75	.01	1.0	-.06
	B	0	18178	14	-.67	.00	1.0	-.02
	C	1	50034	39	-.39	.00	.9	.25
	MISSING ***		1462	1#	-.69	.02		-.02
CE_31	D	0	15892	12	-.67	.00	1.0	-.04
	B	0	33146	26	-.63	.00	1.0	-.02
	A	0	38409	30	-.63	.00	1.0	-.01
	C	0	28054	22	-.58	.00	1.1	.04
	E	1	13024	10	-.49	.00	1.1	.03
	MISSING ***		1741	1#	-.61	.02		-.03
CE_32	C	0	22364	17	-.71	.00	1.0	-.06
	A	0	26024	20	-.67	.00	1.0	-.03
	E	0	18137	14	-.64	.00	1.0	-.01
	D	0	31231	24	-.56	.00	1.1	.08
	B	1	30711	24	-.52	.00	1.1	.01
	MISSING ***		1799	1#	-.66	.02		-.03
CE_33	B	0	30644	24	-.68	.00	1.0	-.05
	E	0	10981	9	-.67	.01	1.0	-.02
	C	0	37886	30	-.63	.00	1.0	.00
	A	0	19074	15	-.60	.00	1.1	.03
	D	1	29647	23	-.50	.00	1.1	.04
	MISSING ***		2034	2#	-.66	.02		-.03

Tabela 23 - continuação

CE_34	D	0	17051	13	-.73	.00	.9	-.09
	C	0	36844	29	-.67	.00	1.0	-.07
	B	0	27138	21	-.63	.00	1.0	-.01
	A	0	30762	24	-.53	.00	1.1	.12
	E	1	15931	12	-.47	.00	1.1	.04
	MISSING ***		2540	2#	-.66	.01		-.04
CE_D35	1	1	58332	64	-.70	.00	1.0	-.28
	2	2	9521	11	-.46	.00	.9	.06
	3	3	8968	10	-.37	.00	1.0	.08
	4	4	5036	6	-.26	.01	1.1	.10
	5	5	6456	7	-.14	.00	1.0	.16
	6	6	1469	2	-.02	.01	1.1	.10
	7	7	550	1	.06	.01	1.1	.07
	8	8	150	0	.21	.03	1.1	.05
	9	9	105	0	.31	.04	1.1	.04
	MISSING ***		39679	30#	-.74	.00		-.49
CE_D36	1	1	45887	54	-.73	.00	1.0	-.29
	2	2	8416	10	-.53	.00	1.0	.01
	3	3	9478	11	-.43	.00	1.0	.05
	4	4	4341	5	-.32	.00	.9	.07
	5	5	4161	5	-.25	.00	1.1	.08
	6	6	7735	9	-.16	.00	1.1	.15
	7	7	1350	2	-.03	.01	.9	.09
	8	8	1692	2	.00	.01	1.1	.10
	9	9	1982	2	.13	.01	1.0	.15
	MISSING ***		45224	35#	-.76	.00		-.53
CE_D37	1	1	83175	88	-.60	.00	1.0	-.22
	2	2	9278	10	-.31	.00	.8	.15
	3	3	1025	1	-.12	.01	.7	.09
	4	4	545	1	-.01	.02	.8	.07
	5	5	276	0	.12	.02	.8	.07
	6	6	384	0	.22	.02	.8	.08
	7	7	64	0	.29	.05	1.2	.04
	8	8	63	0	.37	.05	1.3	.04
	9	9	60	0	.57	.05	1.2	.05
	MISSING ***		35396	27#	-.76	.00		-.54
CE_D38	1	1	7345	9	-1.06	.01	1.2	-.17
	2	2	1477	2	-.94	.01	1.0	-.07
	3	3	13041	16	-.77	.00	1.2	-.12
	4	4	2769	3	-.66	.01	1.2	-.03
	5	5	10824	13	-.58	.00	1.0	-.05
	6	6	13877	17	-.58	.00	1.0	-.13
	7	7	3341	4	-.35	.01	.9	.05
	8	8	6651	8	-.30	.00	.9	.07
	9	9	22909	28	-.16	.00	.9	.32
	MISSING ***		48032	37#	-.79	.00		-.52

Tabela 23 - continuação

CE_D39	1	1	32506	40	-.79	.00	1.0	-.30
	2	2	8264	10	-.57	.00	1.1	-.01
	3	3	8974	11	-.49	.00	1.0	.01
	4	4	6558	8	-.39	.00	1.0	.05
	5	5	6494	8	-.30	.00	1.0	.08
	6	6	8295	10	-.22	.00	1.0	.12
	7	7	2772	3	-.14	.01	1.0	.09
	8	8	2723	3	-.07	.01	1.1	.11
	9	9	4002	5	.05	.00	1.0	.18
		MISSING ***	49678	38#	-.77	.00		-.55
CE_D40	1	1	9218	10	-1.01	.01	1.2	-.15
	2	2	8432	9	-.83	.00	1.2	-.09
	3	3	11813	13	-.69	.00	1.2	-.06
	4	4	12580	14	-.57	.00	1.1	-.02
	5	5	12401	13	-.49	.00	1.1	.00
	6	6	24572	27	-.37	.00	1.1	.09
	7	7	6818	7	-.26	.00	1.0	.09
	8	8	3466	4	-.18	.01	1.1	.08
	9	9	2907	3	-.07	.01	1.1	.10
		MISSING ***	38059	29#	-.79	.00		-.48

* Average measure does not ascend with category score/ # Missing % includes all categories. Scored % only of scored categories

Anexo 4 - Itens e curvas referentes à habilidade das pessoas e probabilidade de resposta por categoria nos itens de FG e CE

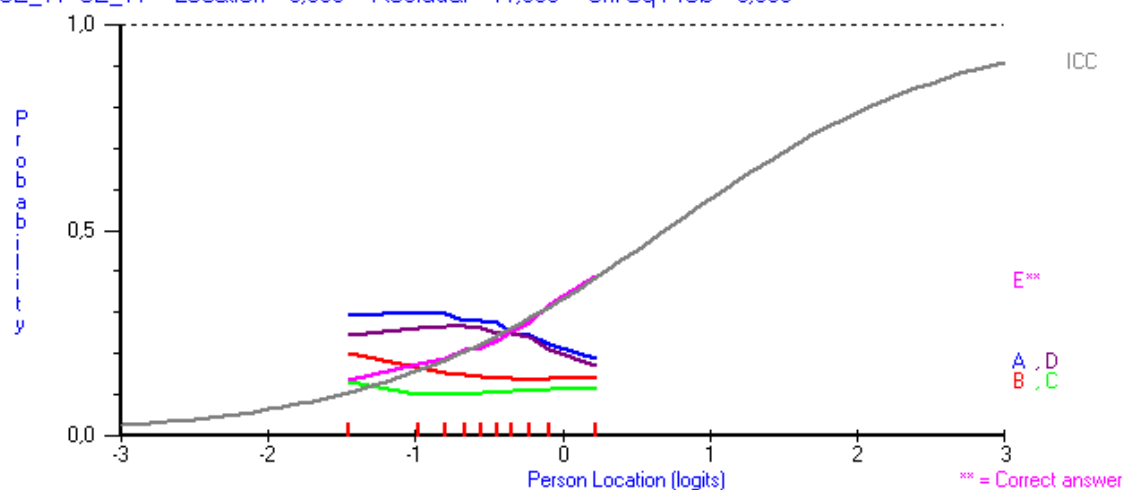
QUESTÃO 11

Desde o início de seu Curso o jovem gerente tinha aprendido que uma das atividades mais desafiadoras do Administrador era tomar decisões, em especial, em grandes empresas. Ele, todavia, estava agora no comando da pequena empresa pertencente à sua família.

Nesta situação, no Brasil, pode-se afirmar que as decisões, na maioria dos casos, tendem a ser

- (A) programadas, ocorrendo raras decisões não programadas.
- (B) concentradas em uma alternativa, pois há limites de tempo.
- (C) condicionadas pela baixa turbulência do ambiente de negócio.
- (D) tomadas com base na racionalidade plena.
- (E) arriscadas, apesar de ser difícil mensurar o risco.

CE_11 CE_11 Location = 0,695 Residual = 11,986 Chi Sq Prob = 0,000



QUESTÃO 12

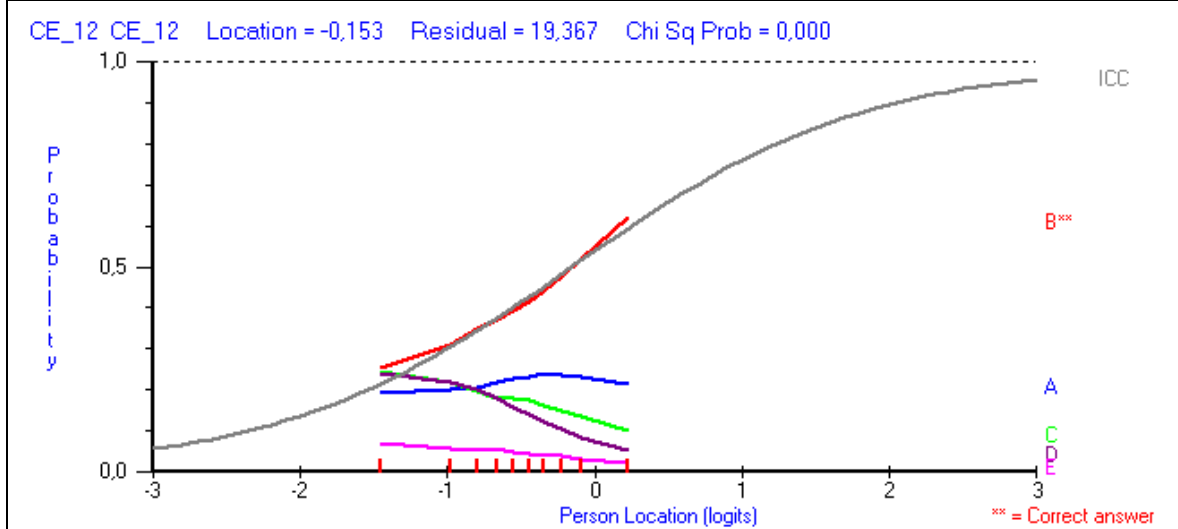
Embora constitua área de conhecimento das mais fascinantes, as bases teóricas da Administração ainda estão em formação. Os estudos pioneiros de Taylor e Fayol, por exemplo, foram ampliados, de forma significativa, nos anos posteriores. Sobre as Teorias da Administração pode-se afirmar que:

- I - na Burocracia, o trabalho realiza-se por meio de funcionários que ocupam cargos, os quais têm atribuições oficiais, fixas e ordenadas por meio de regras, leis ou disposições regimentais;
- II - na Administração Científica, enfatiza-se o estudo das tarefas, a seleção e o treinamento de trabalhadores e a busca pela eficiência operacional;
- III - na Reengenharia de Processos, há um esforço deliberado de se ter uma visão sistêmica da empresa, lastreado em estruturas organizacionais verticalizadas;
- IV - na visão Contingencial, procura-se analisar como as condições ambientais da empresa afetam as possibilidades de escolha nas decisões organizacionais;
- V - na abordagem Comportamentalista, a eficácia organizacional é promovida pela aplicação de análise quantitativa aos problemas e decisões administrativas.

São corretas, apenas, as afirmativas

- (A) I, II e III.
- (B) I, II e IV.
- (C) I, III e V.
- (D) II, IV e V.

(E) III, IV e V.

**QUESTÃO 13**

Na última reunião de direção da Empresa MC – Moderna e Competente, foi analisado o seguinte quadro, elaborado pela Unidade de Inteligência Competitiva:

Ano (*)	Empresa MC			Concorrente Principal			Todo o Setor		
	2004	2005	2006	2004	2005	2006	2004	2005	2006
Participação de Mercado (%)	20	25	30	25	20	10	–	–	–
Margem de Contribuição Média (%)	30	30	30	34	38	42	36	36	36
Idade Média dos Equipamentos (anos)	10	7	5	12	13	14	15	14	13

(*) Dados estimados para 2006.

Levando em consideração estas informações, a direção da MC pode afirmar que

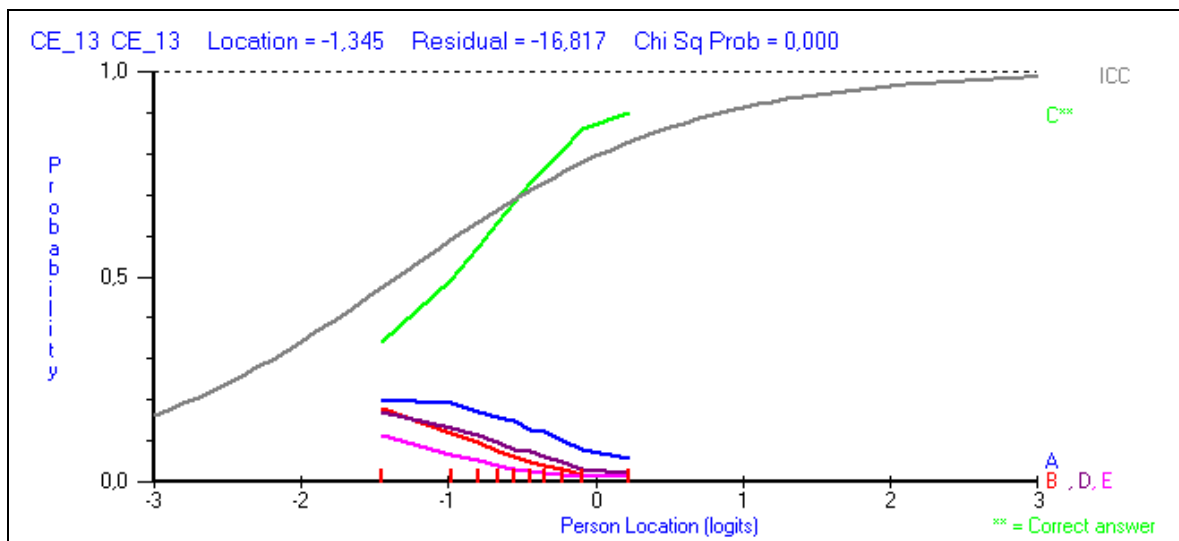
(A) a modernização tecnológica ocorreu de forma mais intensa no concorrente principal, quando comparado com a Empresa MC.

(B) a estratégia competitiva do concorrente principal é ampliar a base de mercado.

(C) a Empresa MC apresenta desempenho mercadológico superior e possui equipamentos mais novos, quando comparada com o principal concorrente.

(D) o setor apresenta instabilidade em termos de rentabilidade, associada a uma lenta modernização tecnológica.

(E) se os custos variáveis, como os de mão-de-obra direta e matérias-primas, são equivalentes nas empresas, os preços da Empresa MC são mais elevados.

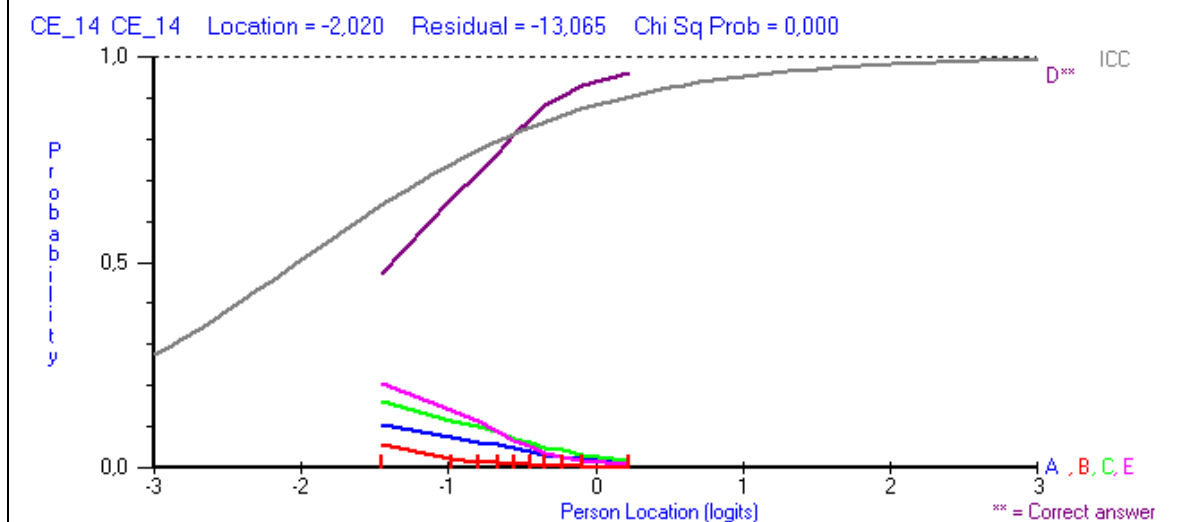


QUESTÃO 14

A Empresa de Exportação Ji-Paraná está fazendo uma atualização de seus cargos e salários. Iniciou o processo a partir da descrição de cargos, como é recomendado.

Para implementar uma descrição eficiente de cargos é necessário considerar

- (A) os requisitos mentais.
- (B) os requisitos físicos.
- (C) as faixas salariais.
- (D) as tarefas e atribuições.
- (E) as condições de trabalho.



QUESTÃO 15

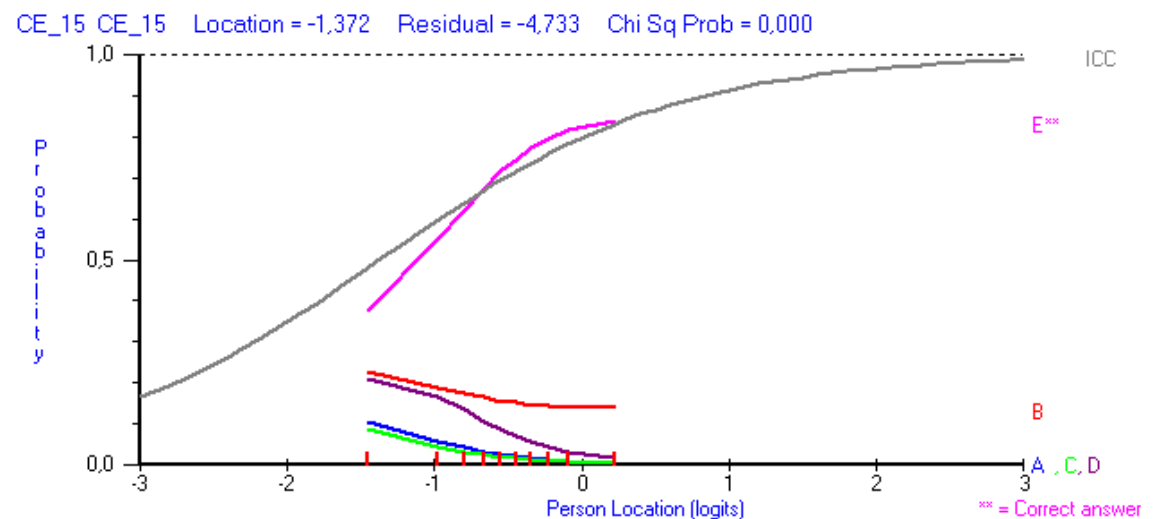
O Banco Solidariedade & Amigos S.A. recém-adquiriu uma empresa de software, a SB Sistemas Bancários Ltda., que detinha 25% do mercado de programas de segurança bancária. Após uma análise organizacional, percebeu-se que a expansão desejada não ocorria devido à falta de agressividade da empresa e da sua estrutura extremamente verticalizada, pesada e inflexível.

O Banco decidiu implementar uma transformação radical na nova empresa, redesenhando-a como uma organização adaptativa. Assim, será necessário desenvolver na empresa uma cultura organizacional que:

- I - seja voltada para o cliente;
- II - valorize a inovação e a criatividade;
- III - mantenha as crenças existentes;
- IV - promova o aprender a aprender;
- V - se baseie em metas e na implantação de um plano de incentivos.

Estão corretos, apenas, os itens

- (A) I, II e III.
- (B) I, II e IV.
- (C) III, IV e V.
- (D) I, II, III e V.
- (E) I, II, IV e V.

**QUESTÃO 16**

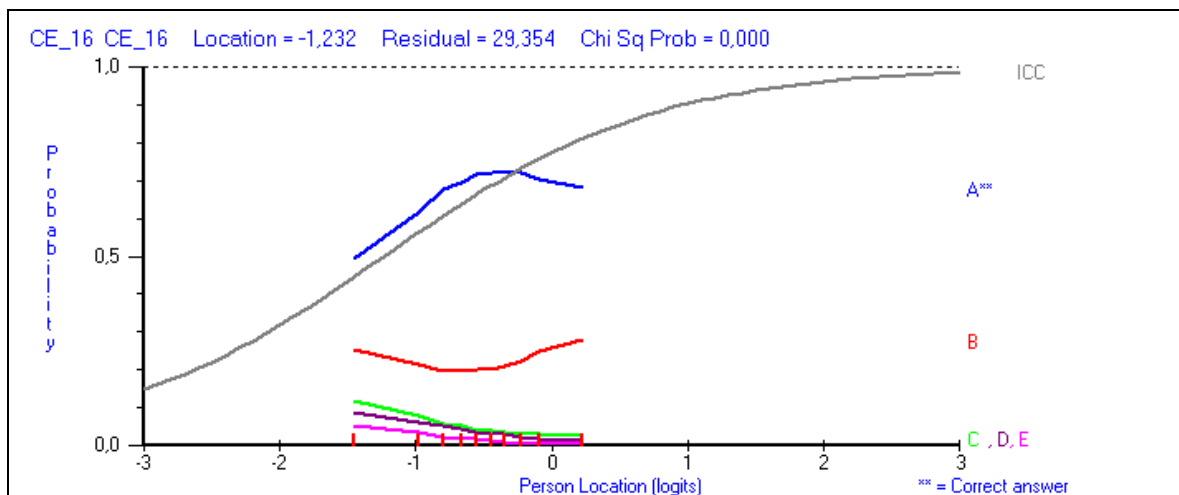
Muitas empresas têm dificuldade de promover mudanças nos comportamentos de seus funcionários no ambiente de trabalho.

PORQUE

As crenças, valores e atitudes que compõem a cultura organizacional influenciam comportamentos dos funcionários na empresa.

Analisando as afirmações acima, conclui-se que

- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.

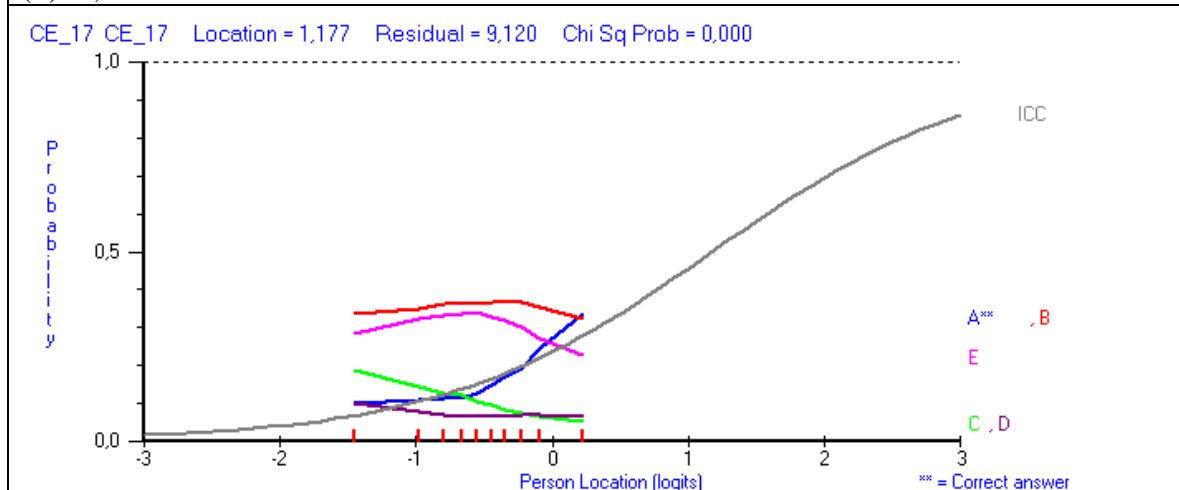
**QUESTÃO 17**

As novas políticas de pessoal, recentemente implantadas na Confecção Norma Modas, trouxeram insatisfação e estresse para o ambiente de trabalho e conseqüentes perdas financeiras. A insatisfação e o estresse foram detectados a partir dos seguintes indicadores:

- I - custos associados a doenças ocupacionais;
- II - violência no trabalho;
- III - nível de responsabilidade atribuída ao cargo;
- IV - inclinações pessoais;
- V - absenteísmo e rotatividade.

Como fontes adequadas de investigação, estão corretos, somente,

- (A) I, II e V.
- (B) I, III e V.
- (C) II, III e IV.
- (D) II, IV e V.
- (E) III, IV e V.

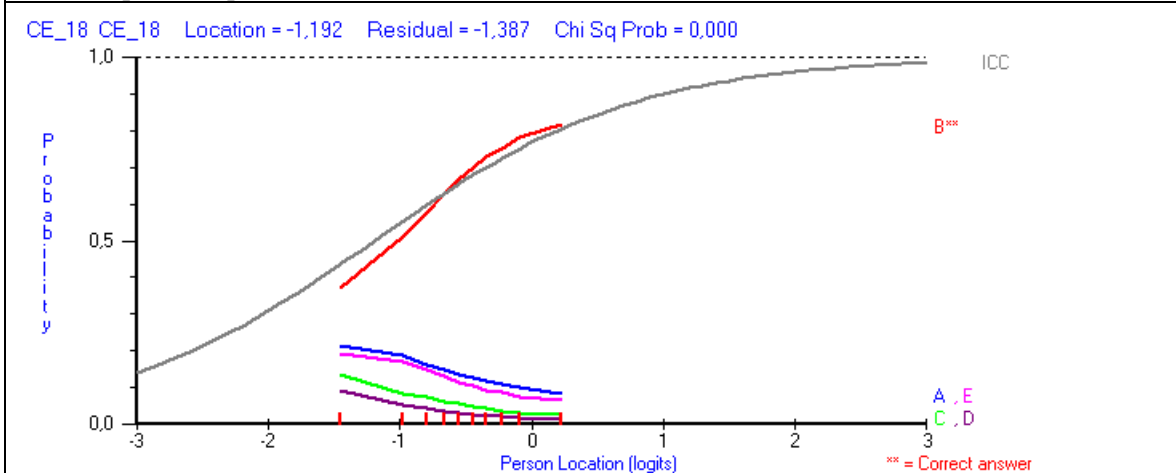


QUESTÃO 18

Dentre os pilotos que compõem o quadro da Cia. Aérea Lunar, alguns têm demonstrado desempenho acima da média. Para esses, o Departamento de Recursos Humanos (RH) resolveu estabelecer um plano de recompensas diferenciado daquele utilizado para o restante da Companhia. Celso, Analista de RH, fez um levantamento das diferentes estratégias para recompensar esse grupo de pilotos.

Quais das estratégias de recompensa relacionadas abaixo são mais indicadas para esse caso?

- (A) Aquelas que estão diretamente vinculadas ao critério dos objetivos de realização empresarial, aliadas ao tempo de serviço no cargo.
- (B) Aquelas que atingem indivíduos de desempenho acima do esperado, sem que o tempo de serviço seja levado em conta.
- (C) Aquelas que contemplam resultados globais, perceptíveis porém impossíveis de serem quantificados.
- (D) Aquelas que contemplam resultados setoriais, perceptíveis porém impossíveis de serem quantificados.
- (E) Aquelas que se referem aos indivíduos de desempenho acima do esperado e que trabalham há mais tempo na empresa.

**QUESTÃO 19**

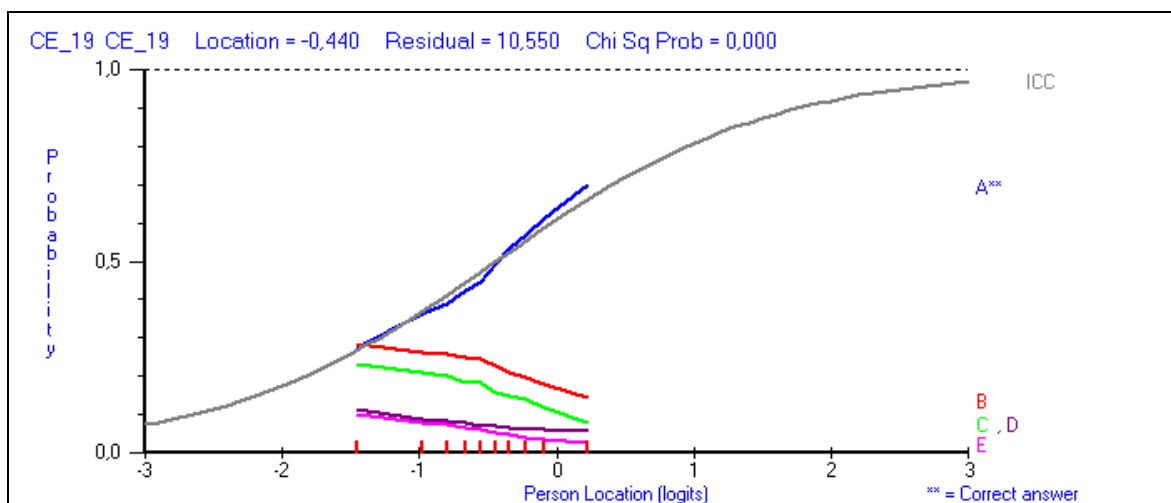
“A Iluminada” é uma empresa que produz luminárias. Ela atende a clientes individuais, vendendo desde luminárias para classes populares até produtos de luxo. A respeito das ofertas da empresa ao mercado, analise as afirmativas abaixo. Ao adotar segmentação de mercado, a empresa pode fixar preços diferenciados de acordo com os segmentos considerados.

PORQUE

As características dos clientes em cada segmento de mercado afetam o valor que eles percebem para os produtos da empresa.

Analisando as afirmações acima, conclui-se que

- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.

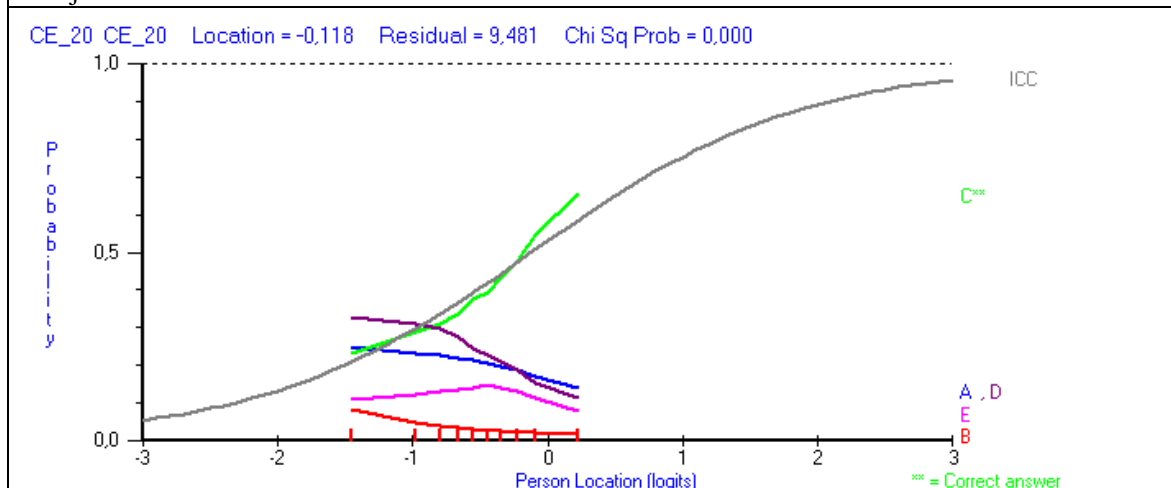


QUESTÃO 20

Jeremias trabalha como Diretor de Marketing da “Pé no Chão”, empresa que produz sandálias de couro. A empresa nunca desenvolveu uma marca de produtor, comercializando seus produtos, no Brasil e no exterior, através de varejistas, que os vendem com suas próprias marcas. Jeremias tem tentado convencer os donos da empresa a desenvolver uma marca de produtor, com a qual eles poderiam vender os produtos.

O que o uso de uma marca de produtor permitiria à empresa “Pé no Chão”?

- (A) Vender seus produtos através de distribuidores exclusivos, o que garantiria maiores margens de lucro.
- (B) Ficar protegida em caso de flutuações na demanda.
- (C) Construir uma imagem junto ao seu público-alvo.
- (D) Posicionar melhor o produto como de alta qualidade combinando uma marca de produtor com preço popular.
- (E) Segmentar o mercado demograficamente, o que não seria possível com marcas próprias de varejistas.



QUESTÃO 21

A “Quitutes da Tia Zuzu” é uma pequena empresa brasileira que produz misturas pré-preparadas para bolos e tortas. O sucesso desta empresa é tal que os diretores estão considerando expandir o composto de produtos, comercializando bolos e tortas prontos. Os gerentes de produto, no entanto, temem que as novas linhas de produtos canibalizem a linha de misturas. Como a empresa pode se proteger da possibilidade de canibalização?

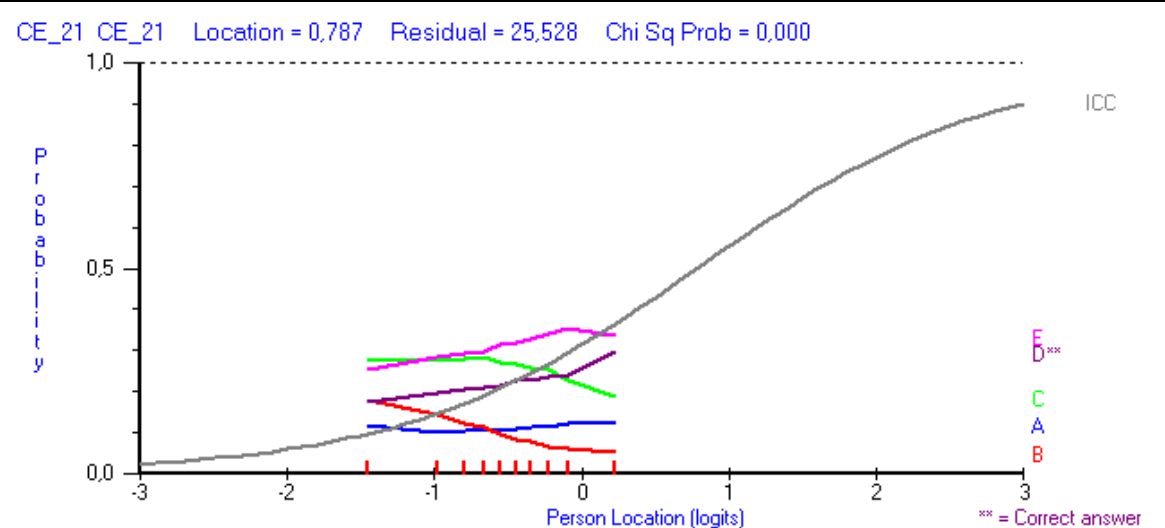
I - Direcionando a comunicação do novo produto para um público com perfil psicográfico diferente daqueles que compram as misturas pré-preparadas.

II - Concentrando a distribuição da nova linha de produtos em regiões onde as misturas de bolos e tortas não são distribuídas.

III - Comercializando a nova linha de produtos com uma marca diferente, usando uma estratégia de multimarcas.

Está(ão) correta(s), apenas, o(s) item(ns)

- (A) I.
- (B) II.
- (C) III.
- (D) I e II.
- (E) I e III.

**QUESTÃO 22**

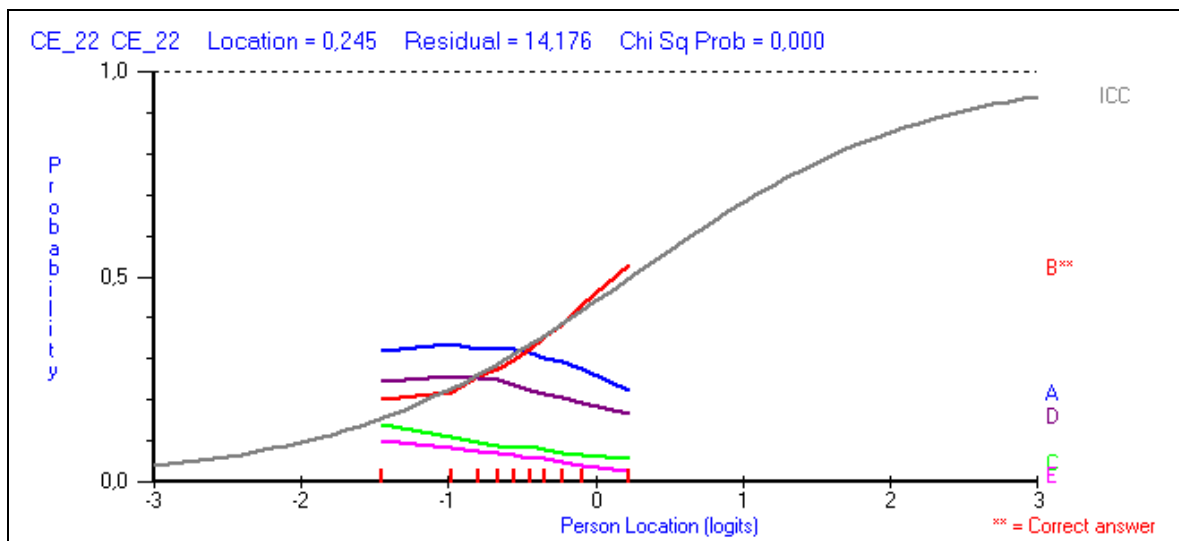
Muitas empresas brasileiras têm tentado exportar os seus produtos. No entanto, a exportação pode requerer que estes sejam adaptados. Sobre isso, analise as afirmativas abaixo. Produtos importados sempre devem ser analisados à luz da cultura local.

PORQUE

A decisão de compra é influenciada, entre outros fatores, pelo nível de conhecimento do consumidor a respeito do produto que compra.

Analisando as afirmações acima, conclui-se que

- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.

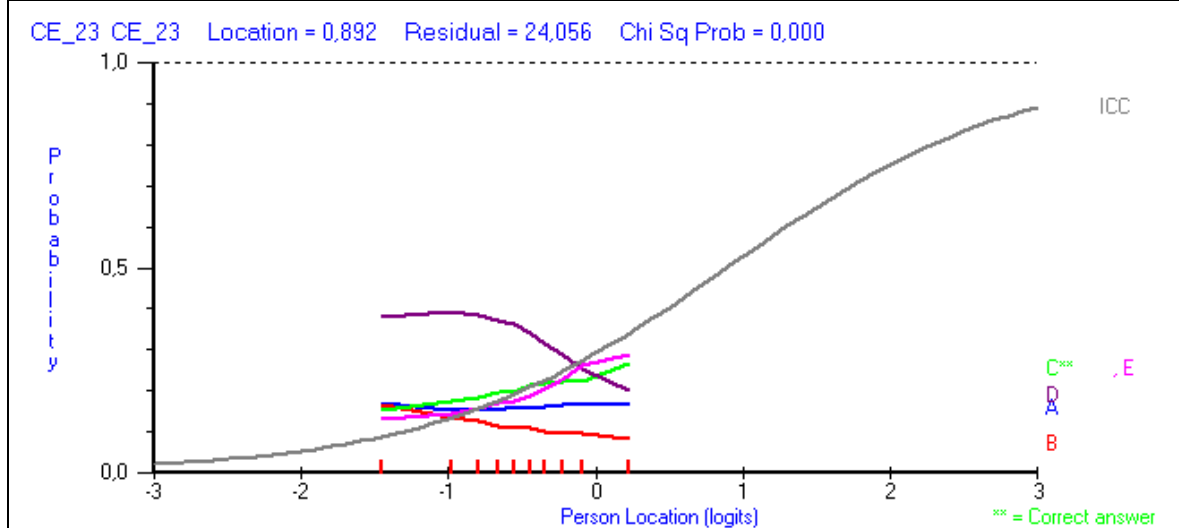


QUESTÃO 23

O novo presidente da Empresa Mercados Múltiplos está redirecionando a sua estratégia por meio de um processo clássico centralizado de gestão estratégica de natureza prescritiva, cujos modelos conceituais e etapas são bem conhecidos.

Assim, o presidente está correto ao afirmar que o processo deve contemplar

- (A) o atendimento das expectativas dos diferentes grupos na estrutura de poder da empresa.
- (B) a implementação de estratégias emergentes.
- (C) a avaliação de concorrentes potenciais e produtos substitutos.
- (D) a utilização de processo participativo de tomada de decisão estratégica.
- (E) a manutenção do organograma e do sistema de informação existentes.



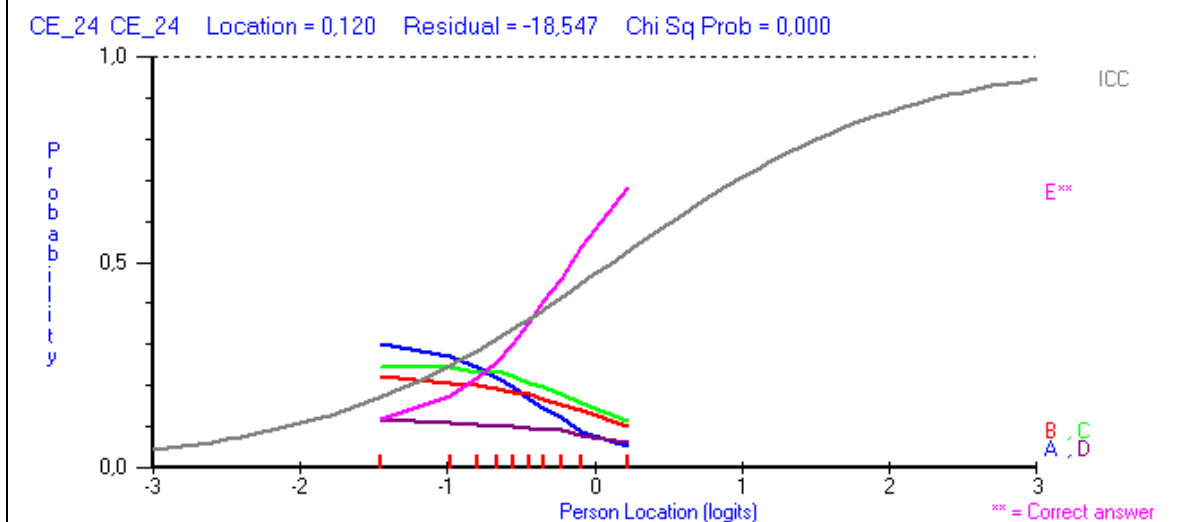
QUESTÃO 24

A Estelar Aérea S.A., uma das mais importantes empresas na vida empresarial recente do Brasil, viveu, durante o ano de 2006, um processo de insolvência, como se pode observar pelos seguintes grupos de contas em seu Balanço de 31 de dezembro de 2005:

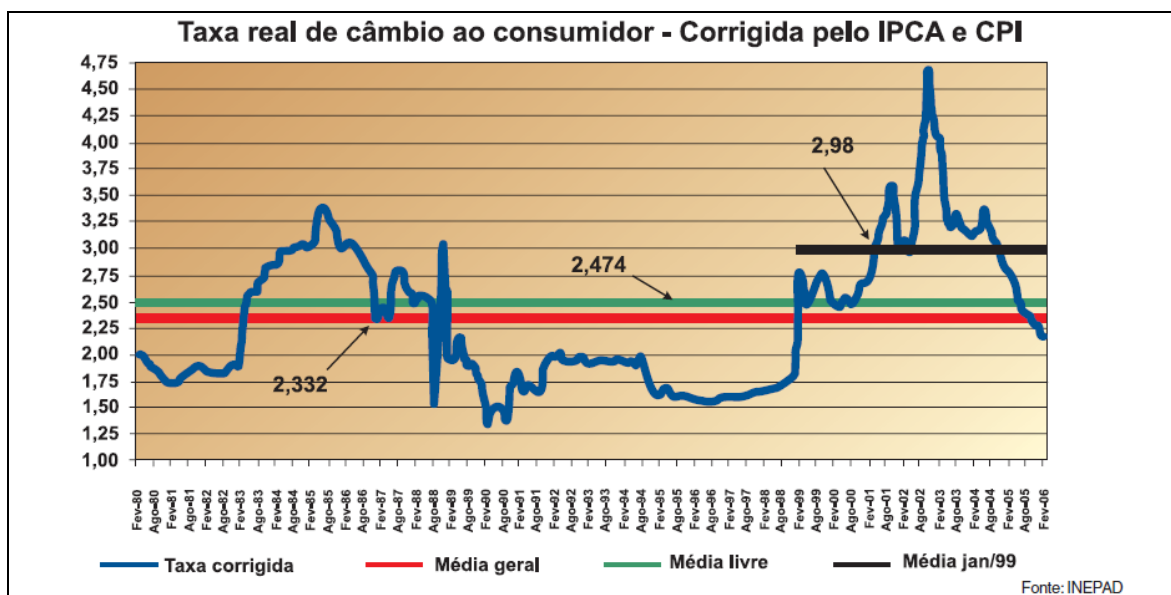
em R\$ milhões	
Ativo Circulante	1.310
Realizável a Longo Prazo	788
Ativo Permanente	238
Passivo Circulante	2.111
Exigível a Longo Prazo	8.146

O Patrimônio Líquido da Estelar Aérea S.A, em 31 de dezembro de 2005, em R\$ milhões, era de

- (A) 12.593
- (B) 7.921
- (C) 2.336
- (D) - 2.336
- (E) - 7.921

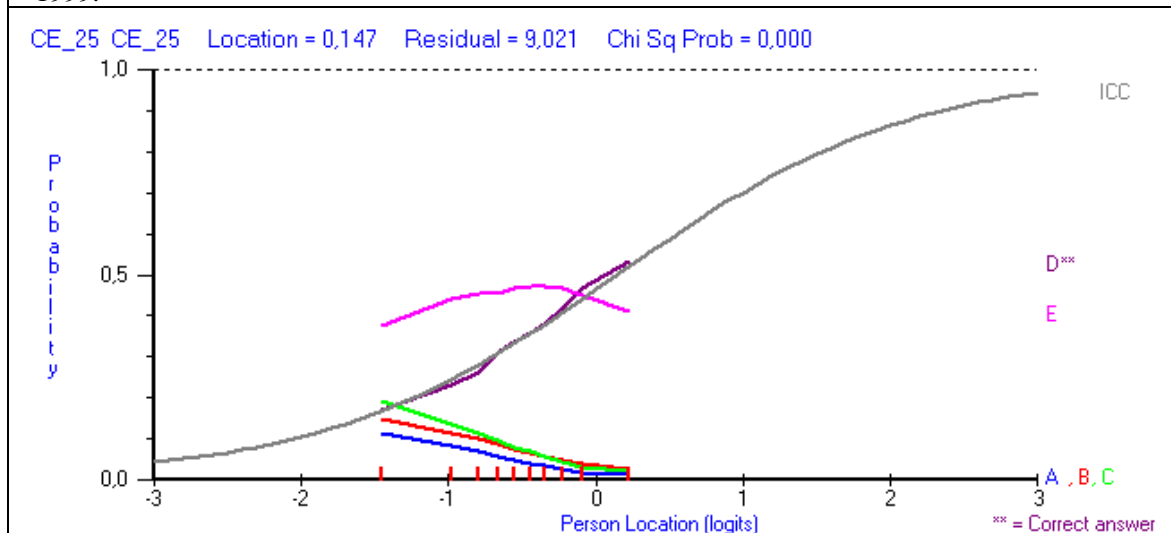
**QUESTÃO 25**

Em razão da valorização do Real diante do Dólar, fomentada por superávits comerciais consecutivos, muitos produtores no Brasil estão preocupados com a impossibilidade de manutenção de suas exportações em 2006 e 2007, passando a rever os volumes a serem produzidos. O gráfico abaixo apresenta a evolução do câmbio real no Brasil versus o Dólar, considerando-se que neles a taxa é corrigida pelo IPCA, Índice de Preços ao Consumidor, e deduzido o CPI, Índice de Preço ao Consumidor americano.



Como se pode observar pelo gráfico, existiram diversos ciclos de câmbio no Brasil no período de 1980 a 2006, com cotações diferenciadas. O ciclo de maior estabilidade cambial foi o período de

- (A) 1982 a 1986, no qual os brasileiros assistiram ao ajuste da economia e ao crescimento do preço do petróleo importado.
- (B) 1986 a 1989, no qual foi implantado o Plano Cruzado, com busca de estabilidade inflacionária e cambial.
- (C) 1990 a 1991, no qual foi implantado o Plano Collor, também com busca de estabilidade inflacionária e cambial.
- (D) 1992 a 1994, no qual o Brasil viveu um período sem planos econômicos.
- (E) 1994 em diante, em razão da implantação do Plano Real, com controle da taxa de câmbio até 1999.



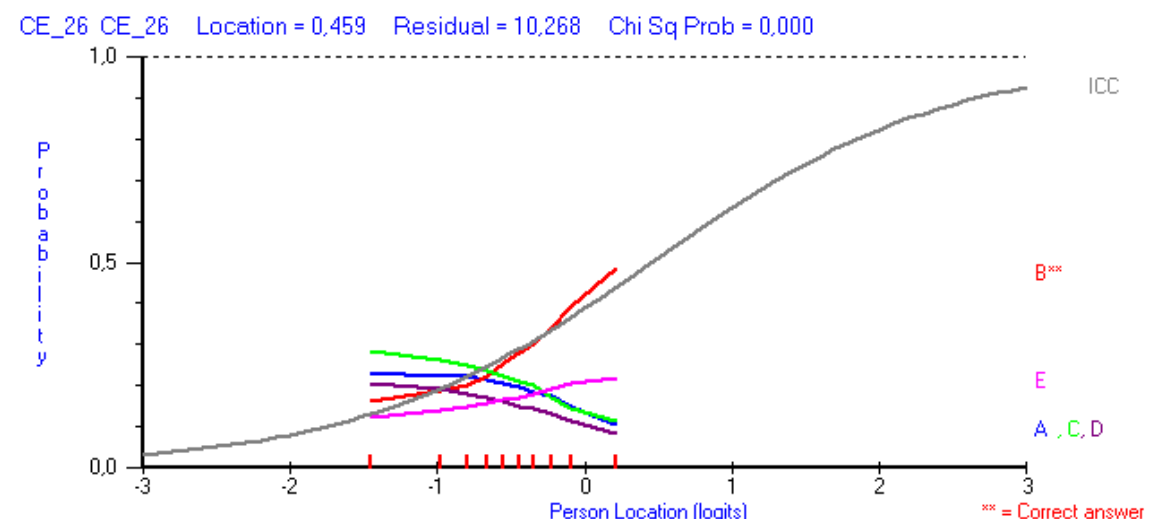
QUESTÃO 26

A CAR (Companhia Agropecuária Rondonópolis) está analisando a implantação de um projeto de investimento no nordeste brasileiro para a produção de frutas com destino ao mercado europeu. Metade do capital necessário ao investimento virá de uma linha de crédito a ser obtida junto ao BNB (Banco do Nordeste do Brasil), e a outra metade virá de capital próprio, a ser captado através do lançamento de ações da empresa no mercado de capitais brasileiro. O investimento será de R\$ 120 milhões, com benefícios anuais líquidos de R\$ 20 milhões, em perpetuidade. O custo de capital junto ao Banco deverá ser de 8% a.a. e o custo do capital próprio é de 12% a.a.. Neste projeto não se deve considerar o Imposto de Renda. A equipe que realizou a análise de viabilidade financeira do projeto encontrou diferentes valores, tais como:

- I - valor presente líquido do projeto: R\$ 80 milhões;
- II - valor presente líquido do projeto: R\$ 200 milhões;
- III - custo médio ponderado de capital do projeto: 8% a.a.;
- IV - custo médio ponderado de capital do projeto: 10% a.a.;
- V - custo médio ponderado de capital do projeto: 12% a.a..

Para se definir a viabilidade financeira do projeto, devem ser utilizados, apenas,

- (A) I e III.
- (B) I e IV.
- (C) I e V.
- (D) II e III.
- (E) II e IV.

**QUESTÃO 27**

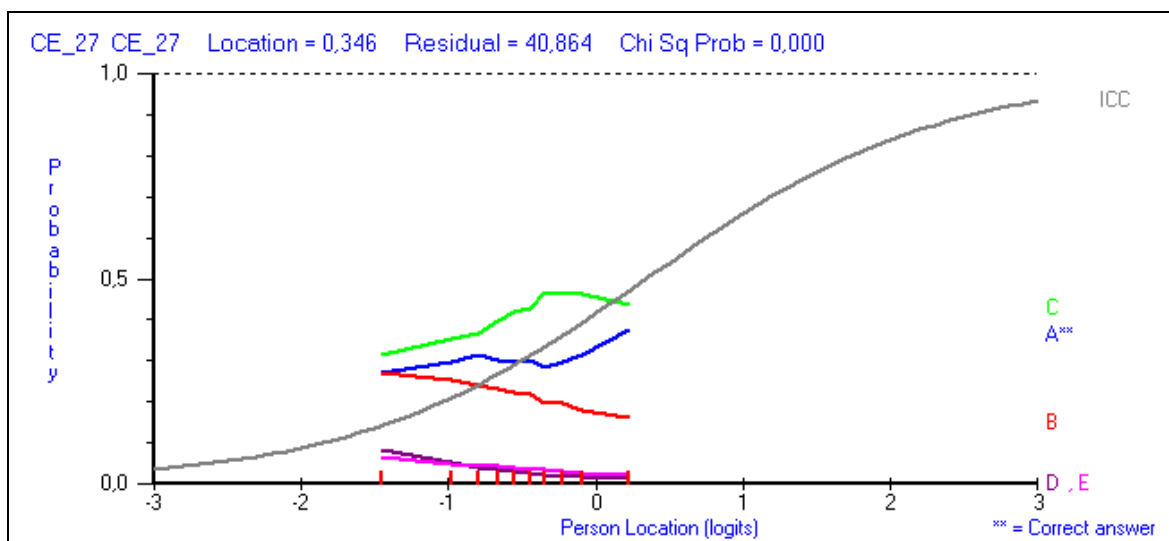
Na maioria dos serviços, o pessoal de atendimento ao cliente necessita de treinamento em habilidades interpessoais.

PORQUE

Na maioria dos serviços, a produção e a entrega acontecem simultaneamente entre o pessoal de atendimento e o cliente.

Analisando as afirmações acima, conclui-se que

- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.



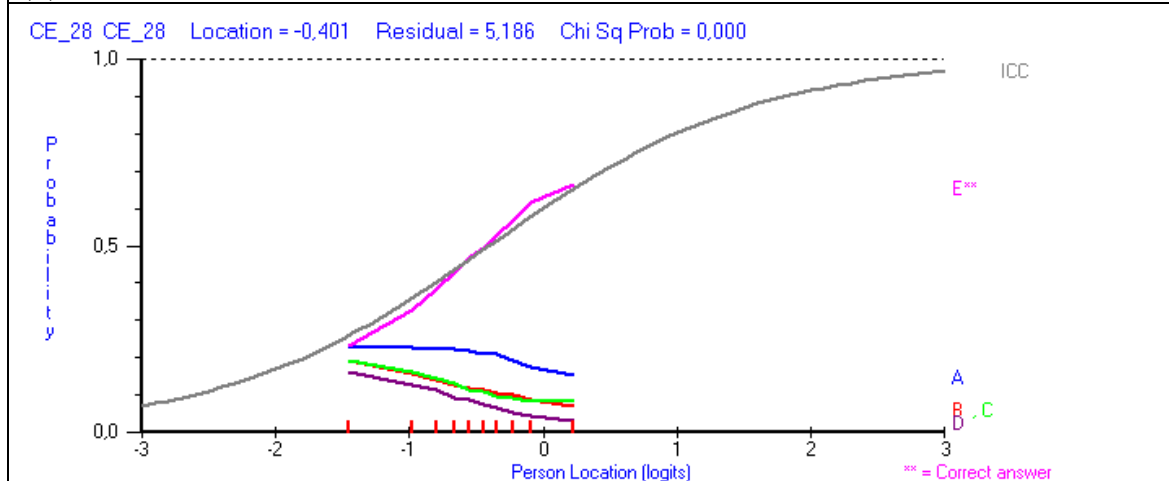
QUESTÃO 28

A seleção de prioridades de implantação de sistemas é uma parte importante do Planejamento Estratégico de Tecnologia de Informação (PETI). Uma visão corporativa das necessidades de sistemas deve ser desenvolvida de tal maneira que as prioridades sejam definidas. Os sistemas implantados em primeiro lugar devem ser aqueles que atendam aos fatores críticos de sucesso do negócio, resolvam problemas imediatos, tenham rápido retorno de investimento ou sejam de implantação rápida ou simples. Constituem atividades do PETI:

- I - definição de estratégias do negócio: diretrizes, planos, objetivos, fatores críticos de sucesso, benefícios do projeto;
- II - início da elaboração do PETI pela definição dos projetos operacionais a serem executados;
- III - verificação da situação atual dos sistemas a serem definidos, avaliando o grau de atendimento das necessidades de informações gerenciais;
- IV - especificação dos equipamentos como base para definição do planejamento;
- V - determinação dos recursos de informática que serão utilizados: conectividades, compatibilidades de hardware, softwares e configuração dos equipamentos.

Estão corretas, apenas, as atividades

- (A) I e III.
- (B) I e V.
- (C) II e III.
- (D) II e IV.
- (E) I, III e V.



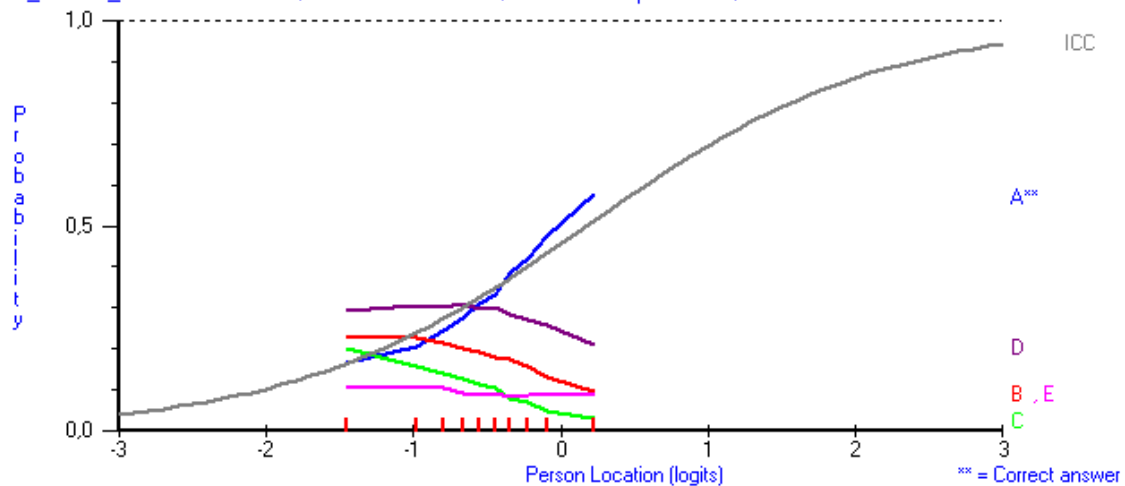
QUESTÃO 29

O atual desafio do administrador da área de Tecnologia da Informação (TI) é projetar e gerenciar a tecnologia para ajudar a empresa a construir vantagem competitiva.

Nesse sentido, pode-se afirmar que Gestão Estratégica de TI consiste em

- (A) identificar oportunidades de diferenciação da empresa no mercado possibilitadas pela TI.
- (B) focar na operação dos sistemas de maneira a permitir sua maior eficiência.
- (C) atualizar a base computacional instalada na empresa de modo a permitir a execução dos processos administrativos.
- (D) habilitar a empresa a alcançar seus objetivos pelo uso eficiente dos recursos disponíveis.
- (E) responder às necessidades de informação pelos diversos usuários da empresa.

CE_29 CE_29 Location = 0,179 Residual = 2,751 Chi Sq Prob = 0,000

**QUESTÃO 30**

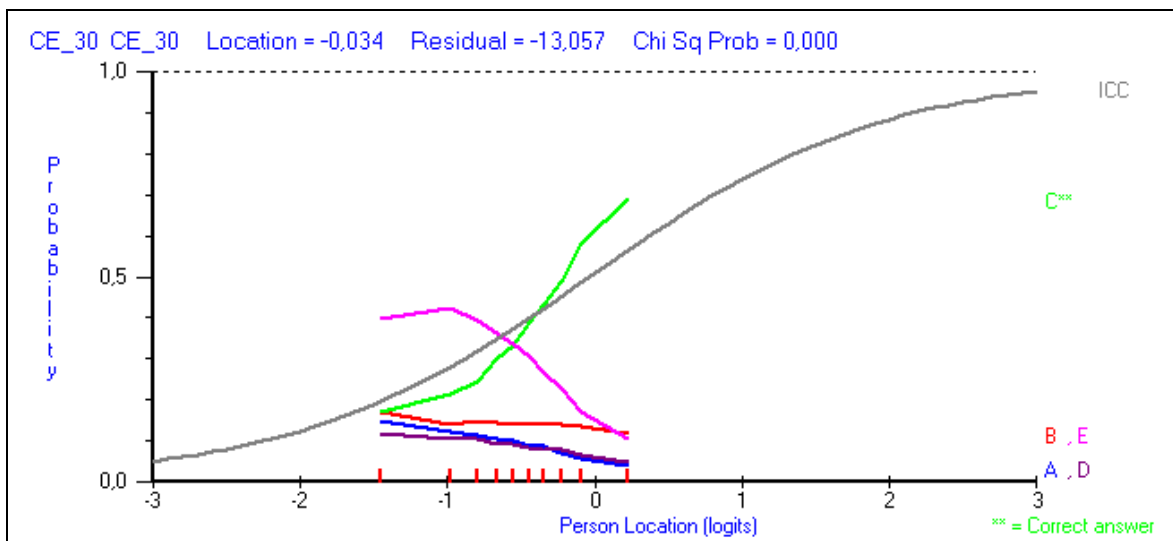
A Cia. Alonso de Auto Peças Ltda. distribui peças para oficinas de reparo de automóveis localizadas em grande área metropolitana. Embora se trate de um mercado competitivo, a Cia. Alonso gostaria de oferecer níveis de estoque adequados às oficinas atendidas, ao mesmo tempo em que deseja maximizar seus lucros. Ela é sabedora de que, à medida que aumenta a percentagem média de atendimentos aos clientes (nível de serviço), maior é seu custo de estoques.

A fim de determinar a influência dos níveis de estoque no percentual de atendimento aos clientes, a Alonso fez um levantamento dos principais itens de seu estoque nos últimos seis meses. A seguinte tabela foi preparada:

Percentagem média de atendimento aos clientes	Nível médio mensal de estoque (R\$)	Custos de estoques mensais (R\$)	Receita média de vendas mensais (R\$)
80%	27.500,00	550,00	900,00
85%	30.000,00	600,00	1.200,00
90%	35.000,00	700,00	1.400,00
95%	40.000,00	800,00	1.450,00
98%	50.000,00	1.000,00	1.600,00

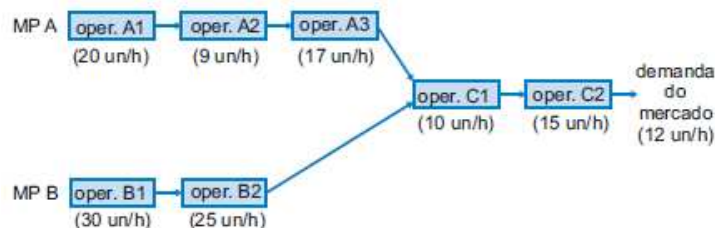
A partir dos dados apresentados nessa tabela, pode-se concluir que o maior lucro ocorrerá quando o nível de serviço for equivalente a

- (A) 80%. (B) 85%. (C) 90%
- (D) 95%. (E) 98%.



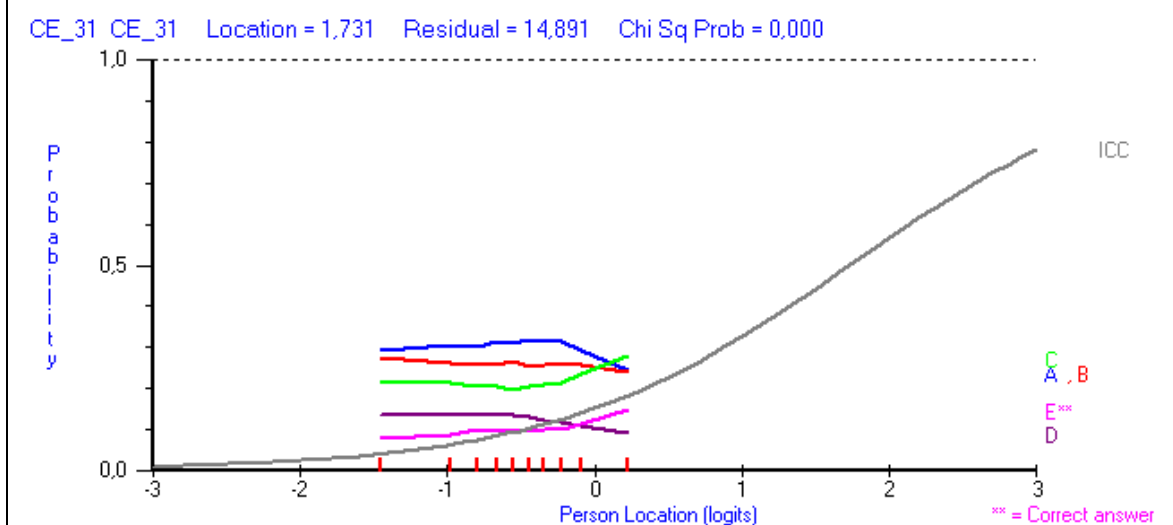
QUESTÃO 31

Observe a figura que se segue, onde a capacidade de produção de cada operação está representada em unidades por hora (un/h). **MP A** representa a matéria-prima A, que recebe seguidamente as operações A1, A2 e A3. **MP B** representa a matéria-prima B, que recebe seguidamente as operações B1 e B2. **C1** representa a operação de montagem dos componentes produzidos a partir de 3 unidades da matéria-prima A e 2 da matéria-prima B. **C2** representa a operação que dá o acabamento final ao produto.



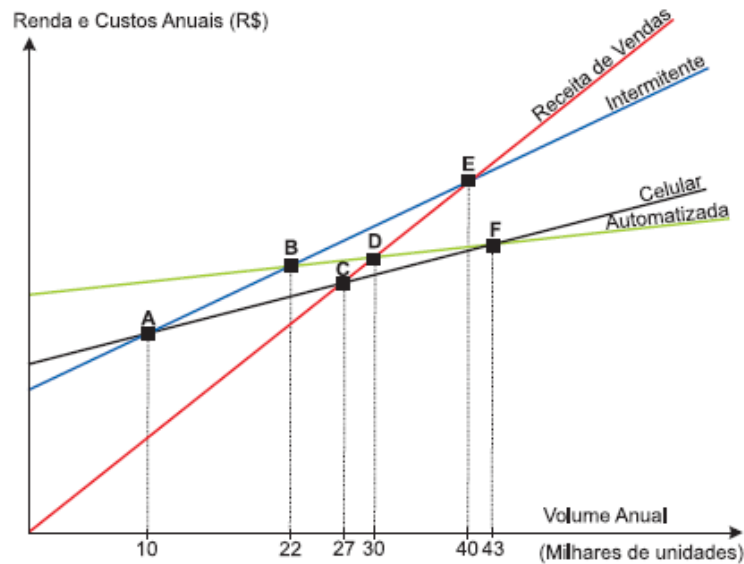
Sabendo que a demanda do mercado é de 12 un/h, de quanto é a produção máxima de produtos acabados?

- (A) 12 un/h
- (B) 10 un/h
- (C) 9 un/h
- (D) 5 un/h
- (E) 3 un/h



QUESTÃO 32

A figura abaixo representa os custos de diferentes formas de processos de produção (celular, automatizada e intermitente), e a receita de vendas de um determinado produto.



Considerando a figura, analise as afirmações a seguir. Se for esperado um volume de produção abaixo de 10.000, a manufatura intermitente é a preferível; entre 10.000 e 43.000, a manufatura celular é a preferível; acima de 43.000, a manufatura automatizada é a preferível.

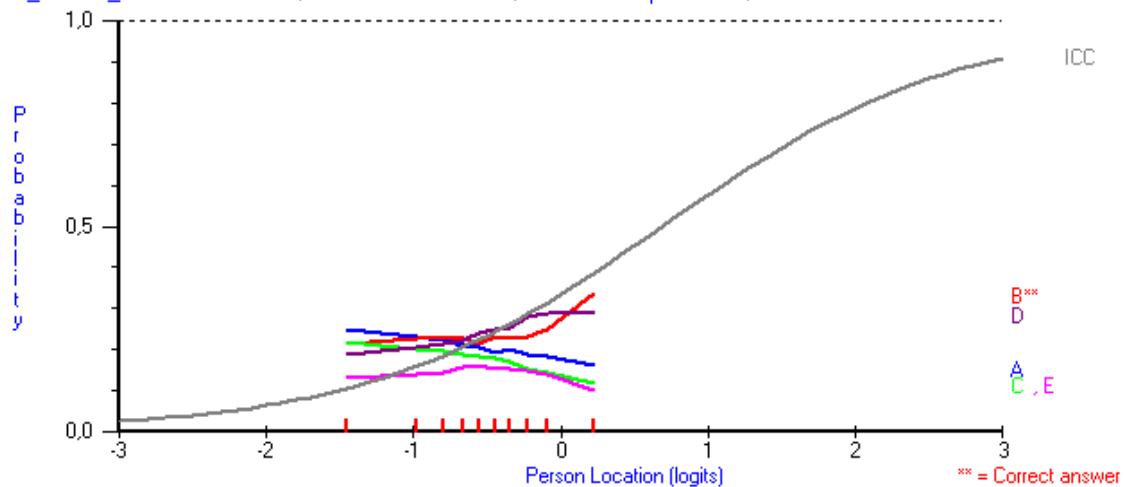
PORQUE

Os pontos de equilíbrio (quantidade/valor para os quais as receitas igualam os custos) são de 27.000, 30.000 e 40.000, respectivamente, para as manufaturas celular, automatizada e intermitente.

A respeito das informações acima, conclui-se que

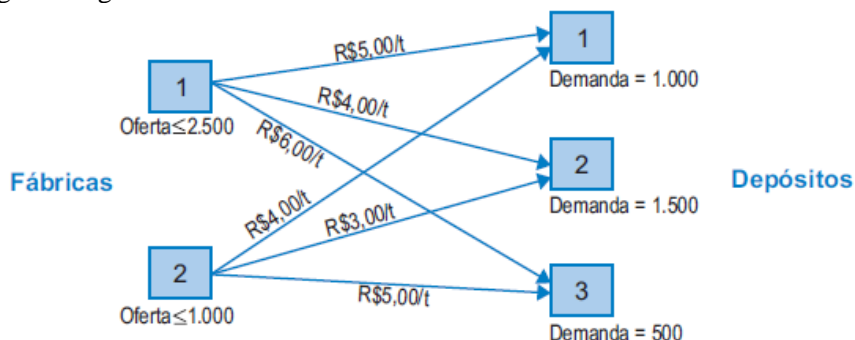
- (A) as duas afirmações são verdadeiras, e a segunda justifica a primeira.
- (B) as duas afirmações são verdadeiras, e a segunda não justifica a primeira.
- (C) a primeira afirmação é verdadeira, e a segunda é falsa.
- (D) a primeira afirmação é falsa, e a segunda é verdadeira.
- (E) as duas afirmações são falsas.

CE_32 CE_32 Location = 0,692 Residual = 31,308 Chi Sq Prob = 0,000



QUESTÃO 33

Analise a figura a seguir.



A Cia. de Produtos Vegetais – CPV possui duas fábricas que abastecem três depósitos. As fábricas têm um nível máximo de produção baseado nas suas dimensões e nas safras previstas. Os custos em R\$/t estão anotados em cada rota (ligação entre as fábricas e depósitos). José de Almeida, estudante de Administração, foi contratado pelo Departamento de Logística com a finalidade de atender a demanda dos depósitos sem exceder a capacidade das fábricas, minimizando o custo total do transporte.

Em sua decisão ele considerou as seguintes situações:

I - 1.000 unidades devem ser transportadas da Fábrica 2 para o Depósito 1. A demanda restante deve ser suprida a partir da Fábrica 1;

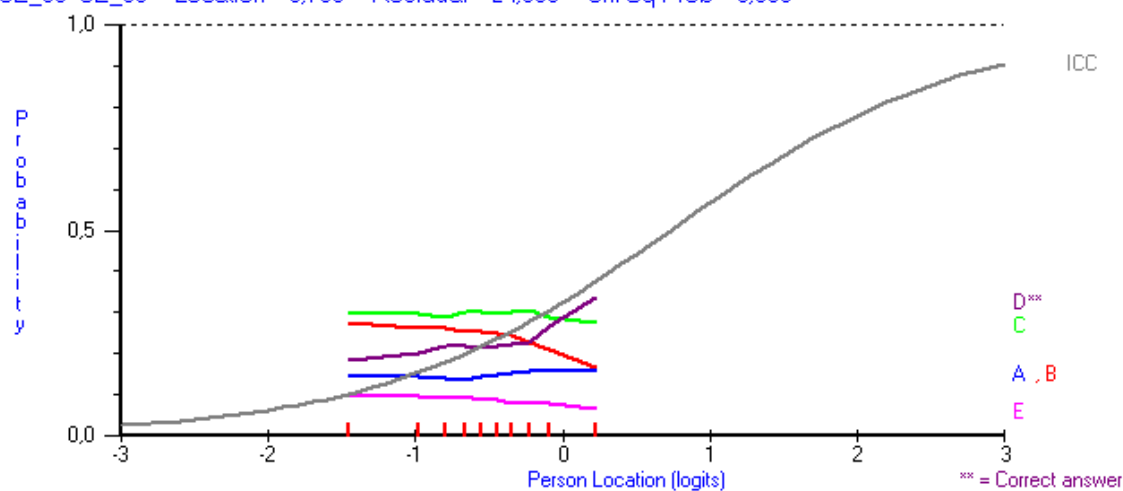
II - 2.500 unidades devem ser transportadas da Fábrica 1 para os Depósitos 1 e 2. A demanda restante deve ser suprida a partir da Fábrica 2;

III - 1.000 unidades devem ser transportadas da Fábrica 2 para o Depósito 2. A demanda restante deve ser suprida a partir da Fábrica 1.

Apresenta(m) o(s) menor(es) custo(s) apenas a(s) situação(ões)

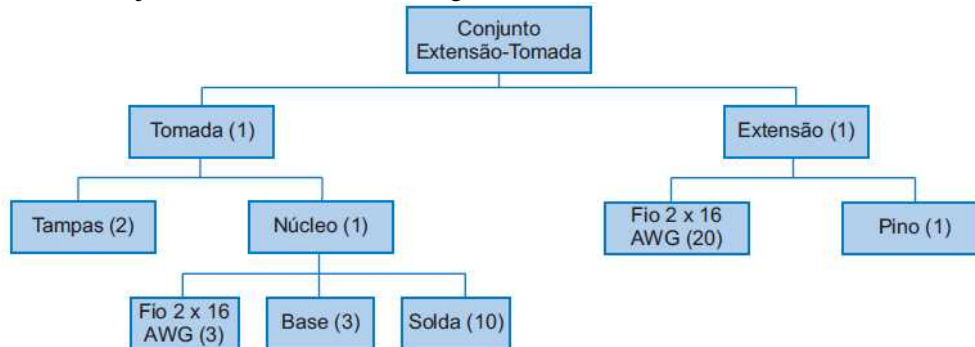
- (A) I. (B) II. (C) III. (D) I e III. (E) II e III.

CE_33 CE_33 Location = 0,735 Residual = 24,536 Chi Sq Prob = 0,000



QUESTÃO 34

A Cia. Goiás Velho S.A., fabricante de conectores, recebeu uma encomenda de 1.200 conjuntos extensão-tomada, cuja árvore de estrutura é a seguinte:

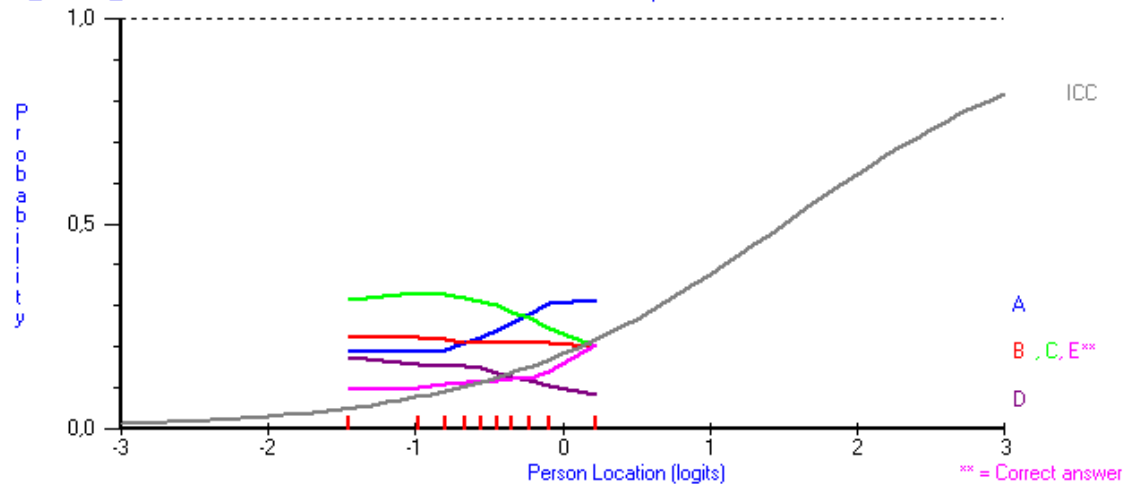


Os números entre parênteses referem-se às quantidades utilizadas na produção de cada conjunto. A Goiás Velho possui em estoque: extensão-tomada = 200; tomada = 100; extensão = 500; fio = 2.000. A nova política de estoques da empresa é a de não manter saldos em estoque, quer em conjuntos, quer em componentes.

A partir das informações apresentadas, pode-se concluir que a quantidade do componente fio (especificação 2 x 16 AWG) que precisa ser adquirido para atender a encomenda de 1.200 conjuntos extensão-tomada (utilizando todo o estoque existente) é

- (A) 25.600
- (B) 21.000
- (C) 12.700
- (D) 11.000
- (E) 10.700

CE_34 CE_34 Location = 1,506 Residual = 16,118 Chi Sq Prob = 0,000



Utilize o texto abaixo para responder às Questões Discursivas de 35 a 38.

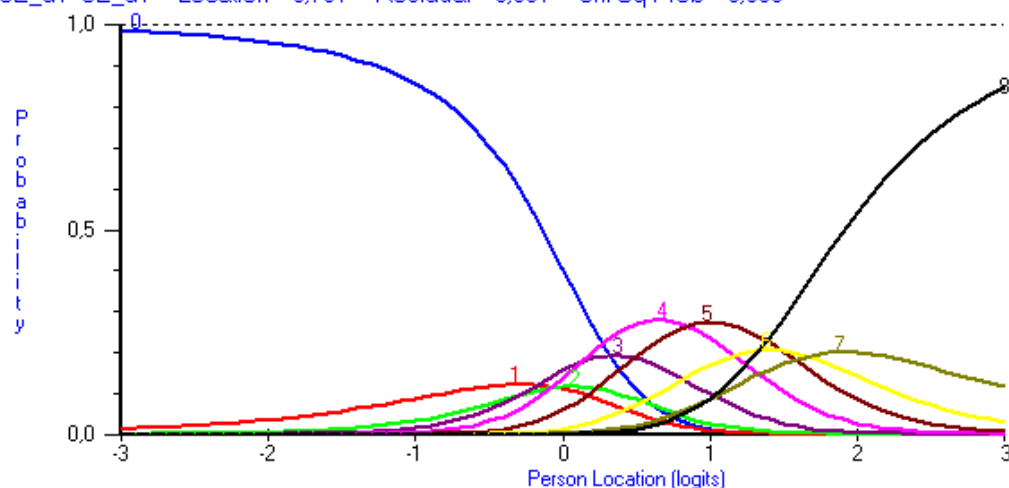
A Companhia Industrial de Tratores está estudando transferir sua atividade de produção de São Paulo para uma unidade maior, buscando aumento de sua produtividade frente à concorrência, para viabilizar sua atividade operacional no Brasil. Atualmente possui duas plantas localizadas no Brasil e seu projeto é integrar a unidade de produção de São Paulo com a unidade de Goiás, ampliando a escala desta última, e vendendo a de São Paulo para empreendimento imobiliário. Estima-se que a paralisação da unidade de São Paulo vá custar R\$ 100 milhões referentes ao programa de desligamento de funcionários, de seus prestadores de serviços e de seus fornecedores de peças. O prédio local e as instalações serão vendidos por esse valor, os quais estão contabilizados a R\$ 50 milhões. Este programa de desligamento envolverá investimentos também em programas de recolocação profissional com reintegração ao mercado de trabalho, apoio psicológico, apoio médico e replanejamento de atividades de prestadores de serviço. Imagine que essas ações aconteçam em 30 dias.

Recentemente, a empresa implantou um sistema integrado de gestão (ERP), disponibilizando informações para o setor financeiro, referentes a faturamento, custos, resultados, orçamento, originando vários relatórios gerenciais. Pela análise desses relatórios, a atual produção de tratores será ampliada em 500 unidades por ano, sem necessidade de investimento adicional em razão de ociosidade na instalação de Goiás. O preço unitário dos tratores será de R\$ 200 mil, e seu custo total unitário, de R\$ 120 mil. A alíquota de Imposto sobre a Renda da Companhia é de 30%. O custo de capital da Companhia é de 10% a.a.

QUESTÃO 35

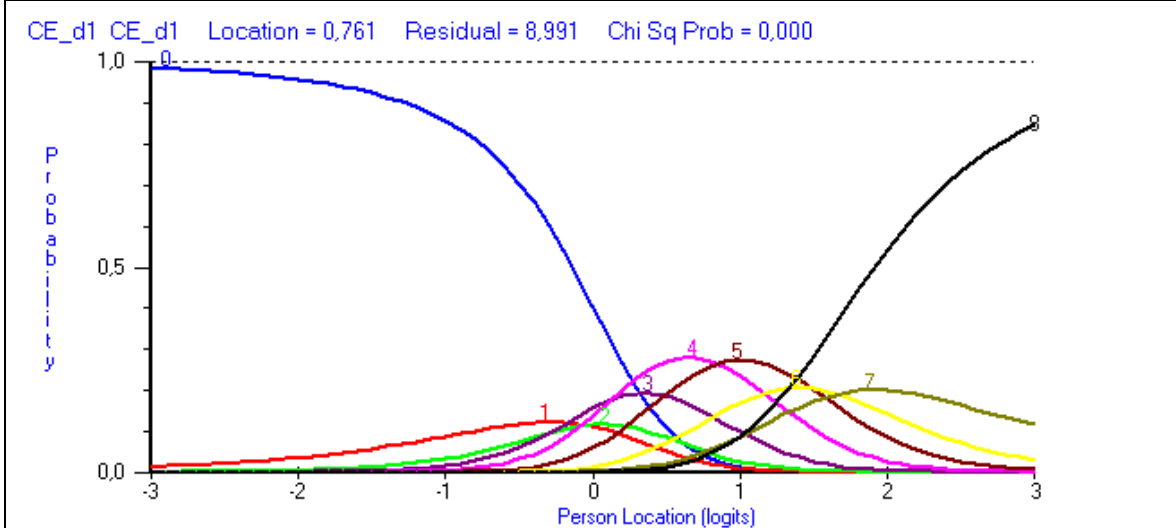
A mudança das instalações da empresa ampliará o ciclo de operações logísticas da Companhia por causa do transporte de tratores de Goiás para São Paulo? Por quê?

CE_d1 CE_d1 Location = 0,761 Residual = 8,991 Chi Sq Prob = 0,000

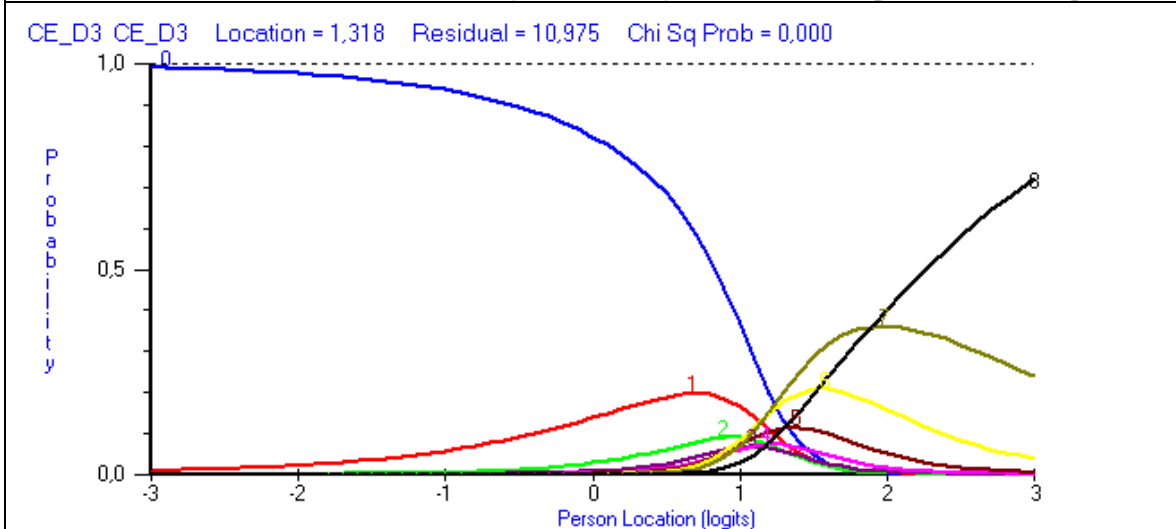


QUESTÃO 36

Cite dois benefícios dos sistemas integrados de gestão (ERP), no suporte à tomada de decisões deste caso, explicando cada um deles.

**QUESTÃO 37**

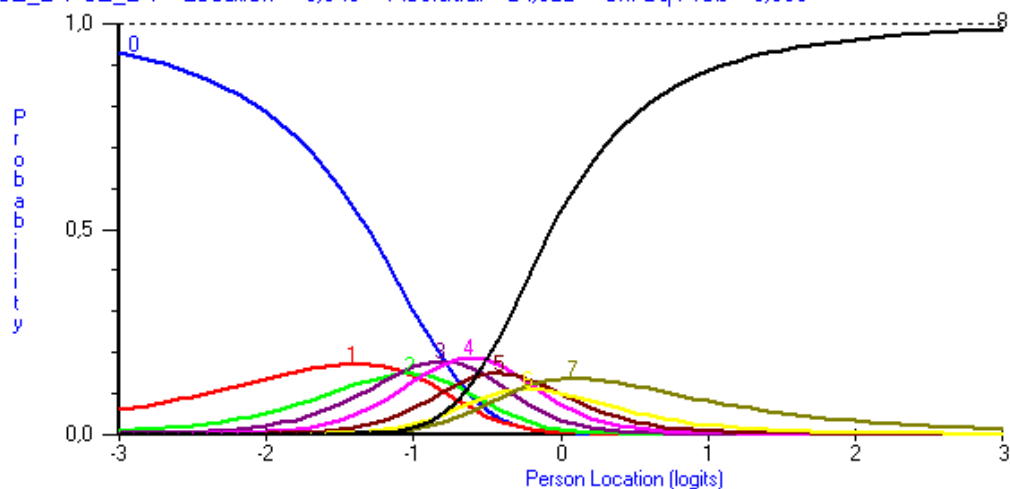
Financeiramente, é recomendável a mudança das instalações de São Paulo para Goiás? Por quê?



QUESTÃO 38

Com base no texto, indique os três grupos com os quais a empresa tem responsabilidade social, neste caso, bem como as respectivas ações que serão realizadas no âmbito da responsabilidade social.

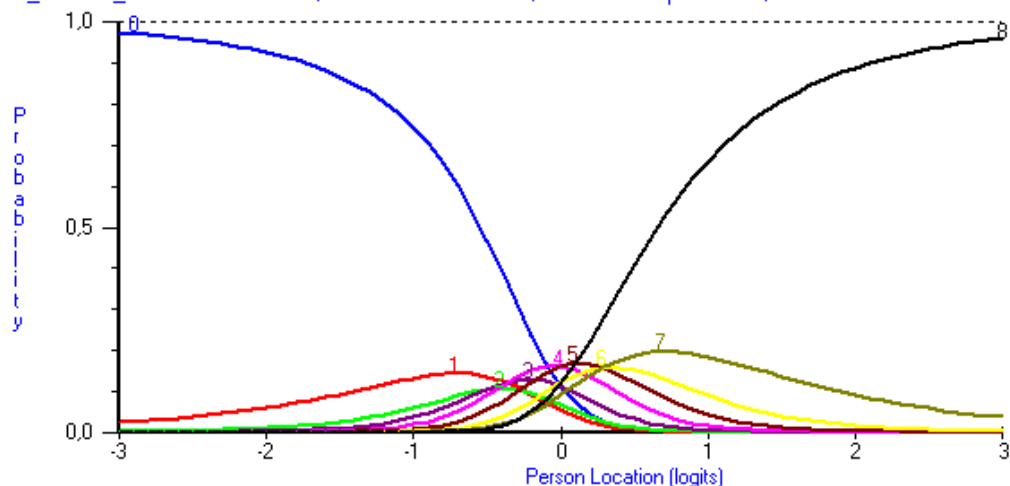
CE_D4 CE_D4 Location = -0,643 Residual = 24,622 Chi Sq Prob = 0,000

**QUESTÃO 39**

A grande imprensa de negócios noticiou recentemente que uma das maiores redes mundiais de fast food estava colocando à venda as lojas próprias que mantinha na América Latina, inclusive no Brasil. Tratava-se de uma mudança importante no modelo de negócio estabelecido, o qual apresentava sinais de esgotamento. Nesse sentido, no novo modelo, a multinacional deixava de ter subsidiárias integrais e passava a receber os royalties pela concessão de sua marca. As lojas seriam vendidas em bloco para grandes franqueadores.

Apresente e justifique uma vantagem e uma desvantagem da franquia como estratégia de expansão de negócios de empresas multinacionais.

CE_D5 CE_D5 Location = -0,006 Residual = 13,658 Chi Sq Prob = 0,000



QUESTÃO 40

Você foi contratado como Gerente da área de Informática da Cia. DEBUG, cujo maior objetivo é a implantação de um novo sistema vital para a empresa, pois a mesma vem perdendo clientes importantes e cometendo erros irreparáveis. Sua equipe já está completa e conta com um grupo de 20 pessoas.

Nesta equipe, existem pessoas que são fundamentais para o bom andamento do processo, pois elas, além de deterem informações, também são respeitadas por todo o grupo. São elas:

- Antônio, que é o mais antigo na equipe, acompanhou todas as implantações anteriores, conhece minuciosamente todos os processos e, em função do seu vasto conhecimento, é respeitado por todos. Entretanto, é resistente a mudanças.
- Sílvia é “um boa praça”. Sempre alegre, adora festas, tudo é motivo de comemoração. Analista de Sistemas altamente técnico, conhece tudo referente a esta implantação e em relação à empresa. Peça-chave no processo, é pouco concentrado e tem estabilidade no emprego, o que o leva a ser um pouco “folgado”.
- Maria, chefe do setor de vendas, é a que mais deverá utilizar o novo sistema. Altamente resistente ao uso de computadores, alega que não precisa do sistema e prefere que aumentem o quadro de pessoal, já que todos os documentos devem ser conferidos manualmente.

Explique como deverá ser a atuação do Gerente em relação a essa equipe, visando a minimizar a resistência à mudança, no processo de implantação do novo sistema.

CE_D6 CE_D6 Location = -0,036 Residual = 80,175 Chi Sq Prob = 0,000

