

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO**

Juliano Anderson Pacheco

**Benchmarking by Item Response Theory (BIRTH): método de
benchmarking utilizando a teoria da resposta ao item na construção
de escalas de competitividade**

Tese submetida ao Programa de
Pós-graduação em Engenharia de
Produção da Universidade Federal
de Santa Catarina para a obtenção
do Grau de Doutor em Engenharia
de Produção

Orientador: Prof. Dalton Francisco
de Andrade, Ph.D.

Coorientador: Prof. Antonio Cezar
Bornia, Dr.

Florianópolis

2014

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

PACHECO, JULIANO ANDERSON

Benchmarking by Item Response Theory (BIRTH) [tese] : método de benchmarking utilizando a teoria da resposta ao item na construção de escalas de competitividade / JULIANO ANDERSON PACHECO; orientador, Dalton Francisco de Andrade; coorientador, Antonio Cezar Bornia. - Florianópolis, SC, 2014. 141 p. ; 21cm

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Produção.

Inclui referências

1. Engenharia de Produção. 2. Benchmarking. 3. Teoria da Resposta ao Item. I. Andrade, Dalton Francisco de. II. Bornia, Antonio Cezar. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Produção. IV. Título.

Juliano Anderson Pacheco

Benchmarking by Item Response Theory (BIRTH): método de benchmarking utilizando a teoria da resposta ao item na construção de escalas de competitividade

Esta Tese foi julgada adequada para obtenção do Título de “Doutor”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção.

Florianópolis, 21 de fevereiro de 2014.

Profa. Lucila Maria de Souza Campos, Dra.
Coordenadora do Curso

Banca Examinadora:

Prof. Dalton Francisco de Andrade, Ph.D.
Orientador
Universidade Federal de Santa Catarina

Prof. Antonio Cezar Bornia, Dr.
Coorientador
Universidade Federal de Santa Catarina

Prof. José Afonso Mazzon, Dr.
Examinador Externo

Prof. Héilton Ribeiro Tavares, Dr.
Examinador Externo

Prof. Rafael Guilherme Burstein
Goldszmidt, Dr.
Examinador Externo

Prof. Carlos Fernando Martins, Dr.
Examinador Externo

Prof. Sérgio Murilo Petri, Dr.
Membro Titular

Prof. Dalton Francisco de
Andrade, Ph.D.
Presidente/Orientador

À minha base e fonte de inspiração e
motivação, Elisiana e Tássio.

AGRADECIMENTOS

Agradeço especialmente ao meu orientador, Prof. Dalton, que acreditou neste projeto e deu-me as condições necessárias para desenvolvê-lo. Além da orientação acadêmica, o mesmo foi um grande educador, pai e amigo.

Ao meu coorientador, Prof. Cezar, que sempre me direcionou para o desenvolvimento de um bom trabalho, puxando “a orelha” quando necessário, com o objetivo único de mostrar o caminho correto.

À minha família, especialmente à minha esposa Elisiana (Nega) e ao meu filho Tássio (Godão), pelo amor incondicional e pela compreensão da minha presença ausente, pois escrever requer dedicação, concentração e, às vezes, esquece-se do nosso entorno.

À Turma da Galinha Pintadinha, aos *Hotwheels* e ao Steve Jobs, pela invenção do *iPod Touch*, que ajudaram na distração e formação do Tássio, proporcionando preciosos minutos de dedicação aos trabalhos.

Aos meus pais, Júlio (*in memorian*) e Marcília, afinal foram eles que começaram realmente este trabalho.

Ao MEC/INEP, em especial ao Carlos Eduardo Moreno Sampaio, Rachel Pereira Rabelo e Sylmara Campos Pinho Garcia, pelo apoio e compartilhamento de informações das pesquisas nacionais sobre a educação superior no Brasil.

As empresas Dígitro Tecnologia e SENAI/SC, quando compreenderam as minhas poucas ausências em prol dos quatro anos deste projeto.

Ao povo brasileiro, que por intermédio dos seus impostos, financiou parcialmente os recursos necessários ao desenvolvimento dos trabalhos.

E se esqueci de alguém, peço desculpas, e agradeço mesmo sem a nomeação formal, todos deveriam saber que foram muito importantes no desenvolvimento deste trabalho!

*Nesta montanha de informação... o objetivo é
fornecer a informação certa, no momento certo,
no formato certo, para a pessoa certa.*

(Armelle Decaup, 2005)

RESUMO

Este trabalho desenvolve um novo método de *benchmarking*, que permite a construção de escalas de competitividade para a comparação de produtos, utilizando-se da Teoria da Resposta ao Item (TRI). Os objetivos foram o desenvolvimento do método e a aplicação demonstrativa em um produto. Teoricamente, o método combina as etapas do processo clássico de *benchmarking* com as etapas da TRI, e demonstra como os modelos estatísticos desta técnica conseguem medir a competitividade de produtos, por meio de um traço latente. Assim, o método da TRI utiliza-se das teorias da Psicometria para mensurar a competitividade de forma qualitativa e quantitativa, pela interpretação das características tangíveis e intangíveis dos produtos. Na aplicação do método desenvolvido, o índice de competitividade será mais assertivo, conforme a aderência do modelo estatístico da TRI às características (itens) do produto, que representam a sua competitividade no mercado. A precisão da medida de competitividade dependerá, basicamente, dos tipos e quantidades de itens elaborados, bem como do número de produtos utilizados na estimativa dos parâmetros do modelo da TRI. Os itens deverão ser construídos com base em teorias pertinentes ao produto e/ou com a opinião de especialistas ou clientes, com o suporte do processo de medição de um traço latente. Para demonstrar a aplicação do método desenvolvido, foram utilizados os microdados do censo da educação superior do Brasil, dos cursos superiores de tecnologia presenciais, totalizando 1.581 cursos. Foram elaborados 16 itens, referentes à dimensão Corpo Docente e Tutorial, provenientes das características das instituições, cursos, docentes e alunos. O referencial teórico do construto foi o instrumento de avaliação de cursos superiores, desenvolvido pelo MEC/INEP. A aplicação do método resulta na construção da escala, nas quais as boas práticas podem ser identificadas pelos itens com maior dificuldade, dado que estes são representados na mesma unidade do índice de competitividade. O acompanhamento ao longo do tempo da competitividade é possível, mesmo com a alteração de características dos produtos, com o uso da equalização da TRI. Assim, os gestores das empresas obtém informações sobre os produtos e o mercado, apoiando-os na avaliação do desempenho frente aos concorrentes, e na tomada de decisão para a melhoria contínua do processo produtivo e ampliação das características de produtos.

Palavras-chave: *Benchmarking*. Teoria da Resposta ao Item. Escala de Competitividade. Cursos Superiores de Tecnologia.

ABSTRACT

This document develops a new method for benchmarking, which allows the construction of scales of competitiveness for the comparison of products using Item Response Theory (IRT). The objectives were the development of the method and its application in a product. Theoretically, the method combines classic benchmarking process steps with IRT steps and demonstrates through statistical models how this technique can measure the competitiveness of products by means of a latent trait. Thus, the IRT method uses the theories of Psychometrics to measure the competitiveness of products through qualitative and quantitative interpretation of the tangible and intangible characteristics of those products. The application of the developed method will increase the accuracy of assessments of the competitiveness of a product because this method uses a statistical model of the IRT to evaluate the characteristics product that reflect market competitiveness. The precision of the measurement of competitiveness will depend basically on the types and quantities of items produced, as well as the number of products used in estimating the model parameters of the IRT. Items will be constructed based on relevant product and/or expert opinion or customers, with the support of the process of measuring of a latent trait. To demonstrate the application of the developed method, the census microdata of higher education in Brazil, the upper courses of classroom technology, totaling 1,581 courses, were used. Sixteen items related to Academic and Tutorial Corps, from the characteristics of institutions, courses, teachers and students were drawn dimension. The theoretical construct of the assessment tool of the superior courses, developed by MEC/INEP. The application of the method results in the construction of a scale in which items identify good practice with greater difficulty because they are represented in the same units as the index of competitiveness. The monitoring along time of the competitiveness is possible, even with the change of product characteristics, using the equalization TRI. Thus, managers of companies obtain information about their products and the market, which allows them to assess their performance against their competitors and to make decisions regarding the continuous improvement of their production process and expansion of product characteristics.

Keywords: Benchmarking. Item Response Theory. Scale Competitiveness. Colleges of Technology.

LISTA DE FIGURAS

Figura 1 - Forças do mercado.....	25
Figura 2 - Estrutura proposta da tese.....	31
Figura 3 - Gerações do <i>Benchmarking</i>	34
Figura 4 - Princípio do processo de <i>benchmarking</i>	36
Figura 5 - Processo referência de <i>benchmarking</i>	37
Figura 6 - Processo para elaboração de uma escala psicométrica.....	49
Figura 7 - Aplicação de Thurstone em 1925.....	50
Figura 8 - Princípio da TRI.....	50
Figura 9 - Aspectos que definem o modelo da TRI.....	52
Figura 10 - Curva Característica de um Item.....	54
Figura 11 - Análise gráfica de curvas características de itens.....	57
Figura 12 - Função de informação do item.....	59
Figura 13 - Função de informação do teste.....	59
Figura 14 - Etapas propostas de uma pesquisa científica.....	61
Figura 15 - Etapas da pesquisa científica para este trabalho.....	63
Figura 16 - Sistematização da revisão bibliográfica.....	65
Figura 17 - Estratégias de buscas no Portal periódicos CAPES.....	66
Figura 18 - Análise das aplicações da TRI.....	67
Figura 19 - Proposta do novo método de <i>benchmarking</i> com base na construção de escalas psicométricas.....	71
Figura 20 - Exemplo de construto da competitividade de um produto, possíveis dimensões e respectivos itens.....	75
Figura 21 - Interpretação da escala de proficiência levando-se em consideração uma característica qualquer que representa um item âncora.....	77
Figura 22 - Representação gráfica do índice de competitividade dos produtos.....	78
Figura 23 - Síntese do método BIRTH.....	80
Figura 24 - Crescimento do número de matrículas em CSTs presenciais e à distância no Brasil na década de 2010.....	85
Figura 25 - SINAES no processo de avaliação do ensino superior no Brasil.....	87
Figura 26 - Curvas características dos itens válidos.....	93
Figura 27 - Curva de informação do teste.....	94
Figura 28 - Escala de competitividade dos CSTs e respectivos itens âncoras.....	95
Figura 29 - Distribuição dos índices de competitividade dos CSTs estimados na escala original pelo modelo ML2.....	97

LISTA DE QUADROS

Quadro 1 - Pontos relevantes nas forças do mercado.....	26
Quadro 2 - Classificação das aplicações de <i>Benchmarking</i>	38
Quadro 3 - Artigos selecionados de aplicações de <i>benchmarking</i>	40
Quadro 4 - Classificação de uma pesquisa científica.	62
Quadro 5 - Equivalência proposta das etapas <i>benchmarking</i> com as etapas de construção de uma escala psicométrica.	70
Quadro 6 - Itens propostos por tipo de variável.	72
Quadro 7 - Exemplos de itens de uma variável binária, de produtos e serviços de informática.....	72
Quadro 8 - Exemplos de itens de uma variável ordinal, de produtos e serviços de informática.....	73
Quadro 9 - Exemplos de itens de uma variável intervalar, de produtos e serviços de informática.....	73
Quadro 10 - Principais iniciativas desenvolvidas pelo MEC/INEP.	84
Quadro 11 - Arquivos constituintes dos microdados do Censo de Educação Superior.....	86
Quadro 12 - Estrutura do instrumento de avaliação dos cursos superiores presencial e a distância.....	88
Quadro 13 - Critérios de atribuição dos respectivos indicadores da dimensão	89
Quadro 14 - Teste da dimensão Corpo Docente e Tutorial e respectivos itens	89

LISTA DE TABELAS

Tabela 1 - A TRI aplicada nas diversas áreas de conhecimento.	67
Tabela 2 - Exemplo de respostas dos cursos aos itens do teste.	90
Tabela 3 - Resultado da análise da distribuição dos itens válidos.....	91
Tabela 4 - Resultado da análise da TRI dos itens válidos.	92
Tabela 5 - Determinação dos itens âncoras.	95

LISTA DE ABREVIATURAS E SIGLAS

AHP - *Analytic Hierarchy Process*
Aneb - Avaliação Nacional da Educação Básica
ANOVA - *Analysis of Variance*
Anresc - Avaliação Nacional do Rendimento Escolar
BIRTH - *Benchmarking by Item Response Theory*
bps - bits por segundo
CCI - Curva Característica do Item
CFA - *Common Factor Analysis*
CI - *Composite Indicators*
Conaes - Comissão Nacional de Avaliação da Educação Superior
CSF - *Critical Success Factors*
CST - Curso Superior de Tecnologia
DEA - *Data Envelopment Analysis*
EAP - *Expected A Posteriori*
EJA - Educação de Jovens e Adultos
Enade - Exame Nacional de Desempenho de Estudantes
Encceja - Exame Nacional de Certificação de Competências de Jovens e Adultos
Enem - Exame Nacional do Ensino Médio
EPT - Educação Profissional e Tecnológica
FII - Função de Informação do Item
FIT - Função de Informação do Teste
GP - *Goal Programming*
HLM - *Hierarchical Linear Modeling*
HOQ - *House Of Quality*
IMMF - *Innovation Management Measurement Framework*
IES - Instituição de Ensino Superior
IIF - *Item Information Function*
INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
IRF - *Item Response Function*
IRT - *Item Response Theory*
ITF - *Test Information Function*
LCD - *Liquid Crystal Display*
LED - *Light-Emitting Diode*
L-LA - *Log-Linear Analysis*
MAP - *Maximum A Posteriori*
MCDA - *Multicriteria Decision Aid*
MCDM - *Multiple-Criteria Decision-Making*

MCP - Modelo de Crédito Parcial - *Partial Credit Model*
MCPG - Modelo de Crédito Parcial Generalizado - *Generalized Partial Credit Model*
MEC - Ministério da Educação e Cultura
MEG - Modelo de Escala Gradual - *Rating Scale Model*
MRG - Modelo de Resposta Gradual - *Graded Response Model*
MRN - Modelo de Resposta Nominal - *Nominal Response Model*
MV - Máxima Verossimilhança
OECD - Organisation for Economic Co-operation and Development
PCA - *Principal Component Analysis*
PISA - Programme for International Student Assessment
QAC - *Quality Award Criteria*
QFD - *Quality Function Deployment*
SAEB - Sistema Nacional de Avaliação da Educação Básica
SEM - *Structural Equation Model*
SENAI - Serviço Nacional de Aprendizagem Industrial
SINAES - Sistema Nacional de Avaliação da Educação Superior
SO - Sistema Operacional
TQM - *Total Quality Management*
TOPSIS - *Technique for Order Preference by Similarity to Ideal Solution*
TRI - Teoria da Resposta ao Item

SUMÁRIO

1 - INTRODUÇÃO	25
1.1 CONTEXTUALIZAÇÃO.....	25
1.2 PROBLEMA DA PESQUISA.....	27
1.3 OBJETIVOS.....	28
1.3.1 Objetivo geral.....	28
1.3.2 Objetivos específicos.....	28
1.4 JUSTIFICATIVA.....	29
1.4.1 Relevância.....	29
1.4.2 Ineditismo.....	30
1.5 DELIMITAÇÕES DO TRABALHO.....	30
1.6 ESTRUTURA PROPOSTA DA TESE.....	31
2 - BENCHMARKING	33
2.1 ORIGEM E CONCEITOS DE BENCHMARKING.....	33
2.2 O PROCESSO DE BENCHMARKING.....	35
2.3 AS CLASSIFICAÇÕES DE BENCHMARKING.....	38
2.4 AS APLICAÇÕES DE BENCHMARKING.....	39
2.4.1 DEA.....	42
2.4.2 AHP.....	45
3 - TEORIA DA RESPOSTA AO ITEM	47
3.1 TEORIA DA MEDIDA.....	47
3.2 ORIGEM E CONCEITOS DA TRI.....	48
3.3 OS MODELOS DA TRI.....	51
3.4 O PROCESSO DE APLICAÇÃO DA TRI.....	56
3.4.1 Calibração dos parâmetros dos itens.....	56
3.4.2 Estimação do traço latente (proficiência).....	60
4 - METODOLOGIA	61
4.1 CARACTERIZAÇÃO DA PESQUISA.....	61
4.2 PROCEDIMENTOS METODOLÓGICOS.....	63
5 - MÉTODO BIRTH	69
5.1 A ADAPTAÇÃO DO PROCESSO.....	69
5.2 O CONSTRUTO, O TESTE E OS ITENS.....	70
5.3 A APLICAÇÃO DOS MODELOS DA TRI.....	75
5.4 A ESCALA DE COMPETITIVIDADE.....	77
5.5 O ÍNDICE DE COMPETITIVIDADE.....	78

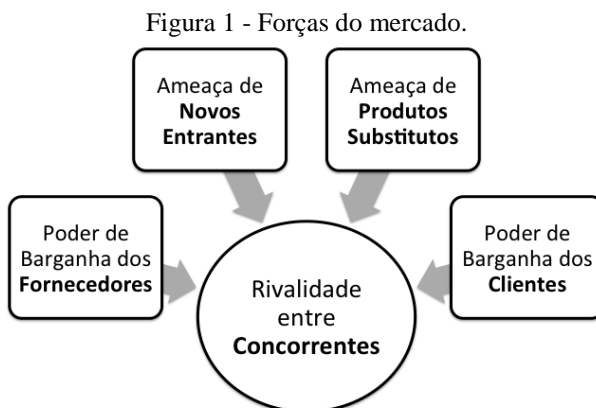
5.6 SÍNTESE DO MÉTODO	79
6 - MÉTODO BIRTH APLICADO NA EDUCAÇÃO SUPERIOR	83
6.1 APRESENTAÇÃO DO PRODUTO EDUCAÇÃO SUPERIOR	83
6.2 O CONSTRUTO, O TESTE E OS ITENS DO ENSINO SUPERIOR	87
6.3 A APLICAÇÃO DO MODELO DA TRI NO ENSINO SUPERIOR	91
6.4 A ESCALA DE COMPETITIVIDADE DO ENSINO SUPERIOR	94
6.5 O ÍNDICE DE COMPETITIVIDADE DO ENSINO SUPERIOR	96
6.6 ANÁLISE DOS RESULTADOS	98
7 - CONCLUSÕES E TRABALHOS FUTUROS	99
7.1 CONCLUSÕES	99
7.2 TRABALHOS FUTUROS	100
REFERÊNCIAS	103
ANEXOS	113
A.1 INSTRUMENTO DE AVALIAÇÃO DE CURSOS DE GRADUAÇÃO PRESENCIAL E A DISTÂNCIA - PARCIAL.....	113
A.2 SCRIPTS DE LEITURA DOS MICRODADOS - IES	123
A.3 SCRIPTS DE LEITURA DOS MICRODADOS - CURSOS.	125
A.4 SCRIPTS DE LEITURA DOS MICRODADOS - DOCENTES	127
A.5 SCRIPTS DE LEITURA DOS MICRODADOS - ALUNOS	129
APÊNDICES	133
A.6 SCRIPTS DA LINGUAGEM R DA APLICAÇÃO DO BIRTH	133
A.7 LOG DA ANÁLISE CLÁSSICA DOS ITENS DA APLICAÇÃO DO BIRTH.....	135
A.8 LOG DAS ESTIMATIVAS DOS PARÂMETROS DOS ITENS DA APLICAÇÃO DO BIRTH	138

1 - INTRODUÇÃO

Este capítulo apresenta a problemática da pesquisa, bem como os objetivos definidos para o desenvolvimento da tese. Apresenta-se a relevância e ineditismo e, também, a estrutura do trabalho desenvolvido.

1.1 CONTEXTUALIZAÇÃO

A dinâmica do mercado, sobre a qual as empresas estão submetidas no longo prazo, é decorrente de forças que regem as relações da cadeia produtiva. Essas forças foram modeladas por Porter (1979) e são: o poder de negociação dos Fornecedores e dos Clientes, a rivalidade entre os atuais Concorrentes, a ameaça de entrada de Novos Entrantes e dos Produtos Substitutos. A Figura 1 apresenta de forma esquemática as forças do mercado.



Fonte: Porter (1979), adaptado.

Em cada força existem pontos relevantes, que influenciam diretamente nas estratégias de atuação das empresas no mercado, que devem ser considerados. No Quadro 1 estão destacados alguns pontos relevantes que essas forças exercem sobre o mercado, segundo Porter (1985).

Neste contexto organizacional modelado por Porter, as organizações, para permanecerem competitivas no atendimento das demandas do mercado, necessariamente precisam aplicar técnicas que permitam acompanhar o estágio de desenvolvimento dos seus produtos,

levando em consideração os insumos disponíveis e os desejos dos clientes, bem como, o nível de competitividade de seus concorrentes diretos e indiretos. Esse acompanhamento tende a gerar conhecimento sobre os produtos e o mercado, a fim de apoiar os gestores na avaliação do desempenho frente aos concorrentes, principalmente, na tomada de decisão para a melhoria do processo produtivo e ampliação das características dos produtos.

Quadro 1 - Pontos relevantes nas forças do mercado.

Forças	Pontos relevantes
Rivalidade entre Concorrentes	<ul style="list-style-type: none"> - número e diversidade de competidores - taxa de crescimento da indústria - barreira de saída - vantagem competitiva sustentável - características dos produtos
Poder de Barganha dos Clientes	<ul style="list-style-type: none"> - grau de dependência com o canal de distribuição - nível de barganha - volume de compra - disponibilidade de informação - disponibilidade de produto substituto - sensibilidade ao preço
Poder de Barganha dos Fornecedores	<ul style="list-style-type: none"> - número de fornecedores - nível de organização do setor - grau de diferenciação dos insumos - custo dos insumos versus preço de venda - ameaça de integração da cadeia produtiva
Ameaça de Novos Entrantes	<ul style="list-style-type: none"> - existência de barreiras de entrada - capital necessário - acesso à distribuição - retaliação esperada dos atuais concorrentes - políticas governamentais
Ameaça de Produtos Substitutos	<ul style="list-style-type: none"> - comprador propenso a substituir - relação custo/benefício - nível percebido de diferenciação

Fonte: Porter (1985), adaptado.

Com a globalização dos mercados, as fronteiras de atuação das empresas são ampliadas e cada vez mais estas enfrentam concorrências nas diversas áreas de negócios, sejam nas atividades de agronegócio, indústria, comércio ou serviços. O conceito de empresa concorrente, segundo Kotler e Keller (2006), está atrelado ao fornecimento de produtos, que suprem as necessidades comuns dos consumidores, ou

seja, a existência de concorrência está intrinsecamente ligada à solução de uma demanda comum de mercado.

O termo produto pode ser definido como um bem material tangível ou serviço intangível produzido por uma empresa, que possui características atreladas ao propósito deste e, geralmente, sintonizadas com as necessidades do mercado consumidor (KOTLER; KELLER, 2006). As empresas que buscam se destacar neste mercado, necessariamente devem desenvolver em seus produtos características diferenciadas que sejam percebidas pelos seus consumidores. A busca pelos diferenciais, as melhores práticas, segundo Camp (1989, p.34), consiste de um “processo de pesquisa coerente à procura de novas ideias, métodos, práticas, processos a fim de adotar as práticas ou adaptar os bons aspectos e implementá-las para se tornar o melhor dos melhores”. Esse processo é denominado *benchmarking* e foi primeiramente utilizado pela Xerox Corporation, no final da década de 1970. O ponto em questão é que a quantificação da competitividade de um produto não pode ser medida diretamente, pois não existe um instrumento e uma unidade de medida para realizar essa medição.

1.2 PROBLEMA DA PESQUISA

O *benchmarking* é uma técnica relevante e aplicável nas empresas, e vem sendo estudada de forma conceitual (FONG; CHENG; HO, 1998, BHUTTA; HUQ, 1999, YASIN, 2002, DATTAKUMAR; JAGADEESH, 2003, ANAND; KODALI, 2008, MORIARTY, 2011) com o foco no entendimento e modificações no processo de aplicação da técnica. São diversas as formas apresentadas para realizar o processo de *benchmarking*, e estas técnicas geram resultados qualitativos e não quantitativos.

Como uma possibilidade de resolver essa lacuna, identifica-se o uso da Teoria da Resposta ao Item (TRI), cujos princípios foram desenvolvidos na década de 1920 (BOCK, 1997), que tem como objetivo a medição de uma variável, que não pode ser observada diretamente, denominada de traço latente (WILSON, 2005). Os modelos estatísticos da TRI permitem medir o traço latente por meio da elaboração de um conjunto de itens relacionados à variável que se quer medir (EMBRETSON; REISE, 2000). Esta técnica é amplamente utilizada em testes psicológicos e em avaliações educacionais. Nas aplicações na área educacional, ela possibilita modelar a probabilidade de apresentar uma certa resposta a um item como função da habilidade

ou proficiência sobre uma determinada área do conhecimento e de características do item respondido (EMBRETSON; REISE, 2000).

Neste contexto, o problema a ser solucionado nesta tese será o desenvolvimento de um novo método de *benchmarking* com o uso de uma técnica quantitativa, que permita identificar quais as empresas (produtos) que possuem as melhores práticas de atendimento ao mercado, mensurando o nível de dificuldade para se obter o mesmo grau de excelência, por intermédio da criação de um índice de competitividade, um traço latente. Para a mensuração desse traço latente será proposto a utilização do modelo estatístico da TRI que melhor represente a competitividade do produto.

Partindo desta problemática, o presente trabalho identifica a seguinte questão:

”COMO A TEORIA DA RESPOSTA AO ITEM PODE CONTRIBUIR PARA AUMENTAR A COMPETITIVIDADE DAS EMPRESAS POR INTERMÉDIO DO BENCHMARKING?”

1.3 OBJETIVOS

Para o desenvolvimento do trabalho definiu-se os seguintes objetivos geral e específicos.

1.3.1 Objetivo geral

O objetivo do trabalho é desenvolver um novo método de *benchmarking* para a construção de escalas de competitividade para a comparação de produtos (bens ou serviços), utilizando-se da técnica quantitativa Teoria da Resposta ao Item (TRI). O novo método foi denominado *Benchmarking by Item Response Theory* (BIRTH).

1.3.2 Objetivos específicos

Para atingir o objetivo geral, definiram-se os seguintes objetivos específicos:

- apresentar a aderência das etapas do processo clássico de *benchmarking* com as etapas da técnica estatística TRI;

- desenvolver um método de *benchmarking* para a mensuração da competitividade de produtos, ao longo do tempo, com o uso dos modelos estatísticos da TRI; e
- desenvolver uma aplicação do novo método de *benchmarking* na construção de uma escala de competitividade de um produto.

1.4 JUSTIFICATIVA

Este trabalho se justifica pela sua relevância e ineditismo apresentados a seguir. A relevância se baseia na importância da técnica *benchmarking* para a competitividade das empresas. A justificativa apresenta como a TRI contribui para a medição da competitividade dessas empresas.

1.4.1 Relevância

Em um ambiente de mercado competitivo, há uma necessidade intrínseca de ferramentas que permitam identificar fatores diferenciais para se atender as necessidades dos consumidores. Esses diferenciais, as boas práticas, as empresas precisam conhecer ou mesmo inserir em seus produtos, fazendo com que esses se modifiquem e que se mantenham competitivos. O acompanhamento, ao longo do tempo, da evolução dos produtos de forma comparativa é também uma característica procurada no monitoramento das melhores práticas de mercado. O *benchmarking* é a ferramenta que visa aumentar o nível de competitividade das empresas, pois permite identificar periodicamente, se a empresa tem as melhores práticas do mercado inseridas nos seus produtos e/ou processos (KOTLER; KELLER, 2006).

Dado que a competitividade não pode ser medida diretamente, por não possuir uma unidade de medida e uma escala de medição, a mesma possui as características essenciais para ser considerada um traço latente (WILSON, 2005). Entende-se que a técnica mais adequada para mensurar um traço latente é a Teoria de Resposta ao Item (TRI) (EMBRETSON; REISE, 2000), em que as boas práticas têm que ser transformadas em itens mensuráveis. Assim, a TRI permite a construção de uma escala de medição e, conseqüentemente, mensura a competitividade de produtos, e também a possibilidade do monitoramento ao longo do tempo dessa competitividade.

1.4.2 Ineditismo

A Teoria da Resposta ao Item (TRI) é uma técnica amplamente utilizada na Psicometria (WILSON, 2005). Foi desenvolvida pela Psicologia (SCHAEFFER, 1998, EMBRETSON; REISE, 2000, COLEMAN *et al.*, 2002) e é também muito aplicada na área educacional (LORD, 1980, ANDRADE; TAVARES; VALLE, 2000). No Brasil, as avaliações do ensino básico, por intermédio do Sistema Nacional de Avaliação da Educação Básica (SAEB), começaram em 1995 a utilizar a TRI para acompanhar a evolução das 4^{as} e 8^{as} séries (atualmente 5º e 9º anos) do ensino fundamental e a 3ª série do ensino médio (SAEB, 2001, KLEIN, 2003). Outras áreas de conhecimento já aplicam a TRI, cita-se a nutrição (WILDE; GERALD; FRIEDMAN, 2004), como, por exemplo, na forma de alimentação, qualidade de vida, na medicina (YESAVAGE; BRINK; ROSE., 1983, WARE; BJORNER; KOSINSKI, 2000, MESBAH; COLE; LEE, 2002; DAS; HAMMER, 2005), bem como nas avaliações do nível de stress, na administração (COSTA, 2001, BAYLEY, 2001, BORTOLOTTI, 2003, BORTOLOTTI, 2010), para a medição da satisfação dos clientes e na própria engenharia de produção (ALEXANDRE *et al.*, 2002a, ALEXANDRE *et al.*, 2002b, TEZZA; BORNIA; ANDRADE, 2011), no tocante à adoção de práticas de qualidade.

Apesar da ampla aplicabilidade da TRI em diversas áreas de conhecimento, não foi encontrado na literatura o uso da TRI como uma técnica quantitativa de desenvolvimento do processo de *benchmarking*. A busca foi realizada de forma sistemática, conforme apresentada na metodologia deste trabalho.

E é nesta lacuna que este trabalho propõe que as etapas de análise da TRI sejam usadas como uma técnica quantitativa e aderente às etapas do processo clássico de *benchmarking*. A proposição é que as características do produto definirão o questionário e, conseqüentemente, os itens e a proficiência possibilitarão a criação da escala de competitividade.

1.5 DELIMITAÇÕES DO TRABALHO

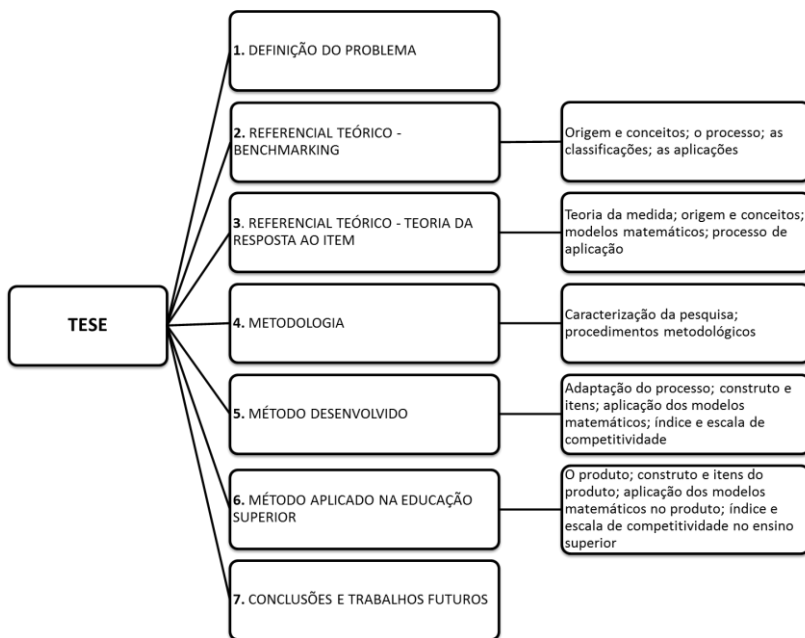
O presente estudo não tem o intuito de modificar a essência do processo clássico de *benchmarking*, mas sim de ser uma opção de método para a aplicação deste.

Destaca-se, também, que a aplicação do método foi desenvolvida para exemplificar a utilização deste e não para exaurir o tema competitividade do produto escolhido. Entende-se que a competitividade de qualquer produto tem relevância para o desenvolvimento de trabalhos específicos sobre o tema.

1.6 ESTRUTURA PROPOSTA DA TESE

A tese está estruturada em 7 capítulos, conforme Figura 2. No capítulo 1 está apresentada a definição do problema. No capítulo 2 é desenvolvido o referencial teórico, onde é explorado o processo de *benchmarking*. No capítulo 3 apresenta-se a Teoria da Resposta ao Item (TRI) como proposta de técnica quantitativa de *benchmarking*. No capítulo 4 apresenta-se a metodologia utilizada no trabalho.

Figura 2 - Estrutura proposta da tese.



Fonte: Elaborado pelo Autor (2011).

A contribuição teórica do trabalho é apresentada no capítulo 5, com a aplicação dos conceitos da Psicometria, em que cada modelo estatístico da TRI possibilita uma forma distinta de mensurar a competitividade de um produto, constituindo como parte integrante do processo de *benchmarking*. No capítulo 6, fez-se o desenvolvimento de uma aplicação exemplo, que consistiu da avaliação de serviços educacionais na modalidade de Educação Profissional e Tecnológica (EPT), um produto intangível. E finalmente, no último capítulo, tem-se as conclusões e a proposição de possíveis trabalhos futuros.

2 - BENCHMARKING

Este capítulo foi desenvolvido para apresentar o tema principal do trabalho: o *benchmarking*. O intuito é apresentar definições pertinentes ao tema, as mudanças da técnica ao longo do tempo, bem como as etapas do processo clássico de *benchmarking*.

Para a demonstração do ineditismo do trabalho, focou-se no levantamento sistemático das técnicas quantitativas utilizadas e, quando houver, das aplicações feitas de *benchmarking* nas diversas áreas de conhecimento.

2.1 ORIGEM E CONCEITOS DE BENCHMARKING

Na prática, o termo *benchmarking* é explorado por vários pesquisadores da área, em que esses, ao seu modo de pensar, definem a influência da aplicação do processo na competitividade das empresas. O que seria *benchmarking* para estes pesquisadores? Uma seleção de possíveis respostas seria:

Benchmarking é:

...o processo contínuo de medição de produtos, serviços e práticas em relação aos mais fortes concorrentes, ou às empresas reconhecidas como líderes em suas indústrias (CAMP, 1989).

...um processo contínuo e sistemático para avaliar produtos, serviços e o trabalho de organizações que são reconhecidas como representantes das melhores práticas com a finalidade de melhoria organizacional (SPENDOLINI, 1992).

...um foco externo nas atividades, funções ou operações internas, a fim de atingir uma melhoria contínua (MCNAIR; LEIBFRIED, 1992).

...a busca contínua por, e aplicação de, práticas significativamente melhores que levam ao desempenho competitivo superior (WATSON, 1993).

...um processo disciplinado que começa com uma busca minuciosa para identificar as melhores

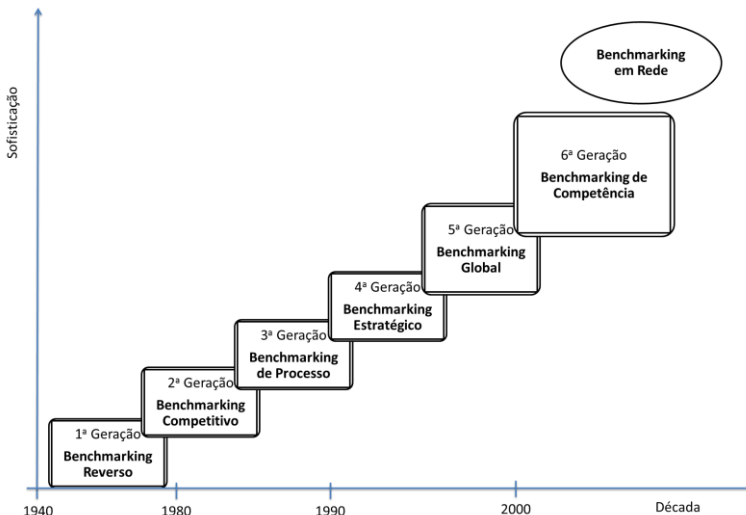
práticas de organizações, continua com o estudo cuidadoso da nossa própria prática e de desempenho, evolui através de visitas sistemáticas e entrevistas, e conclui com uma análise dos resultados, desenvolvimento de recomendações e implementação (GARVIN, 1993).

O que aparece de comum nestas definições? O uso da comparação com referências, como forma de medição e a implantação da melhoria contínua, por meio da sistematização do processo de *benchmarking*.

E como foi a trajetória do processo de *benchmarking* ao longo do tempo? O processo atual de *benchmarking*, como uma prática empresarial, começou formalmente no final da década de 1970, em aplicações primeiramente realizadas pela Xerox Corporation (CAMP, 1989). O *benchmarking*, segundo Watson (1993), originou-se de teorias desenvolvidas desde a década de 1940. Watson foi vice-presidente da Xerox em Soluções da Qualidade e desenvolveu práticas de *benchmarking* na empresa desde 1984 (WATSON; YOUSSEF, 1994).

Na Figura 3 estão apresentadas as gerações do *benchmarking* ao longo do tempo, desenvolvido primeiramente por Watson (1993), sendo adaptado por Ahmed e Rafiq (1998) e Kyrö (2003).

Figura 3 - Gerações do *Benchmarking*.



Fonte: Kyrö (2003), adaptado.

Segundo Ahmed e Rafiq (1998) e Kyrö (2003), cada geração tem um foco específico do uso do *benchmarking*:

- **reverso** (décadas de 1940 a 1970) - engenharia reversa de produtos com foco em características, funcionalidades e desempenho. Considerada a geração que desenvolveu as teorias para o *benchmarking* “moderno”;
- **competitivo** (décadas de 1970 e 1980) - foco nos processos entre competidores de um mesmo setor. É a geração que iniciou a aplicação do *benchmarking* “moderno”;
- **de processo** (décadas de 1980 e 1990) - foco no compartilhamento de informações de processos comuns entre empresas de setores distintos;
- **estratégico** (décadas de 1980 e 1990) - intuito de identificar as melhores estratégias para aumento de desempenho entre potenciais parceiros;
- **global** (décadas de 1990 e 2000) - objetivo de identificar as melhores práticas globais;
- **de competência** (décadas de 1990 em diante) - intuito de promover mudanças comportamentais nas pessoas, promovendo a aprendizagem organizacional;
- **em rede** (década de 2000 em diante) - formar redes para compartilhar boas práticas, independente do tamanho e setor da empresa.

As duas primeiras gerações tem um foco na busca pelas boas práticas nos concorrentes. A terceira e quarta geração utiliza-se de não concorrentes para identificar as melhores práticas. Já as três últimas gerações, buscam todas as possibilidades das melhores práticas, independente do setor de atuação no mercado. Percebe-se que há diferenças importantes entre cada geração, seja na forma de realizar, foco, referência utilizada, abrangência, mas todos tem em comum o uso do processo de *benchmarking* para a identificação das melhores práticas no mercado.

2.2 O PROCESSO DE BENCHMARKING

Benchmarking é, na sua essência, um processo com o intuito de gerar aprendizado com as melhores práticas do mercado, de forma a tornar a empresa mais competitiva.

Assim como as definições (CAMP, 1989; SPENDOLINI, 1992; MCNAIR; LEIBFRIED, 1992; WATSON, 1993; GARVIN, 1993), há algumas variantes na definição do processo de *benchmarking*. Entretanto, as questões levantadas por Watson (1993) definem bem os princípios do processo de *benchmarking*. As questões colocadas por Watson são: o que devemos avaliar? Quem devemos avaliar? Como é que vamos executar o processo? Como é que eles realizam o processo? De forma esquemática, os princípios definidos por Watson (1993) estão representados na Figura 4.

Figura 4 - Princípio do processo de *benchmarking*.



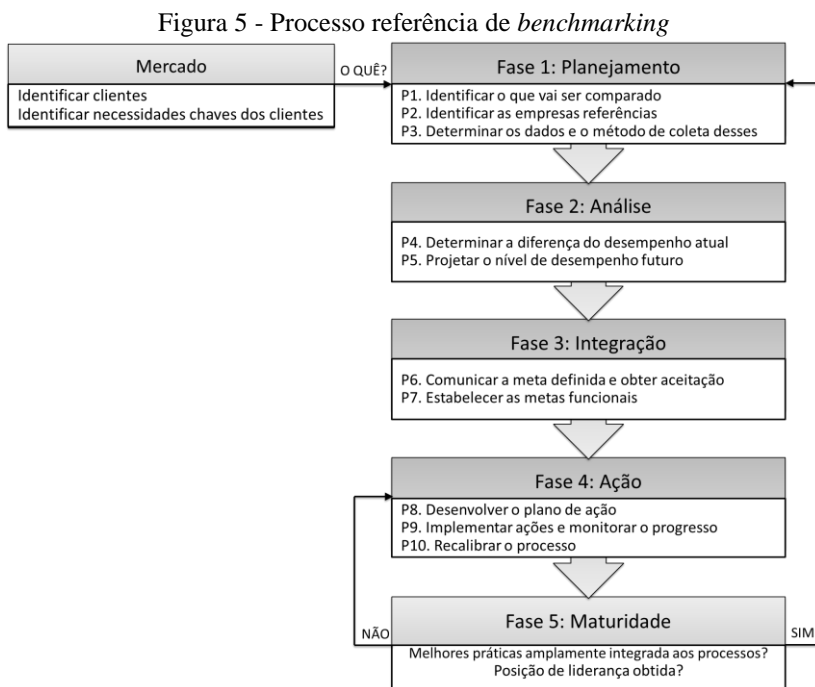
Fonte: Watson (1993), adaptado.

Nota-se pela Figura 4, que “o que” e “de quem” comparar, gera a coleta e a análise de dados de mercado, com o objetivo de verificar como os melhores fazem, para modificar como é feito atualmente. Essa modificação são as saídas, os resultados, para levar as melhores práticas à empresa, seja num produto, processo ou método.

Existem vários modelos do processo de *benchmarking* existentes na literatura (AHMED; RAFIQ, 1998). Entretanto, todos são originários de um único modelo clássico desenvolvido na Xerox e apresentado por Camp (1989). Foi com base neste modelo que surgiram outros, todos variantes do processo da Xerox de *benchmarking* (CARPINETTI; MELO, 2002). Como exemplo, em Anand e Kodali (2008), há uma proposta de 12 etapas, compostas de 54 passos. Entretanto, entende-se que o modelo apresentado por Fong, Cheng e Ho (1998), que alinhou ao modelo precursor da Xerox, uma abordagem sistemática de melhoria do desempenho, de forma a atender as necessidades e exigências dos clientes para promover a satisfação nos mesmos (Figura 5), consiste do processo de referência do *benchmarking* moderno.

Na Figura 5, o processo referência inicia com o direcionamento ao mercado, mediante o atendimento das necessidades destes, sendo o mesmo composto de dez passos, agrupados em cinco fases:

- **planejamento**: definir o objetivo da análise, as empresas referências, o escopo e a técnica de coleta dos dados;
- **análise**: identificar os pontos fortes das empresas de melhores práticas e comparar com o seu desempenho frente aos concorrentes, medindo a diferença e projetando o desempenho desejado;
- **integração**: com base nos resultados dos dados coletados, estabelecer as metas para atingir o desempenho desejado e a comunicação dessas para toda a empresa;
- **ação**: implementar os planos estabelecidos para chegar nas metas, avaliando-os e redefinindo-os periodicamente;
- **maturidade**: medir se a empresa alcançou o desempenho desejado, confirmando a efetiva integração das metas ao processo de gestão da empresa.



Fonte: Fong, Cheng e Ho (1998), adaptado.

Percebe-se que o processo de *benchmarking* é cíclico e realimentado, e que é definido em função dos objetivos que as empresas desejam atingir, sintonizados aos desejos dos clientes, seja na melhoria contínua de processos ou nas características de seus produtos.

2.3 AS CLASSIFICAÇÕES DE BENCHMARKING

As aplicações de *benchmarking*, segundo Fong, Cheng e Ho (1998), são classificadas segundo a natureza da referência, do conteúdo e do propósito de relacionamento. No Quadro 2 estão os tipos e definições relativas às classificações do *benchmarking*.

Quadro 2 - Classificação das aplicações de *Benchmarking*.

Classificação	Tipo	Descrição
Natureza da referência	Interno	comparando dentro de uma organização sobre o desempenho de unidades de negócios ou processos similares
	Concorrente	comparando com os concorrentes diretos, alcançar ou até mesmo superar seu desempenho global
	Indústria	comparando com empresa do mesmo setor, incluindo não concorrentes
	Genérico	comparando com uma organização que se estende além dos limites da indústria
	Global	comparando com uma organização em que a sua localização geográfica se estende além das fronteiras nacionais
Conteúdo de benchmarking	Processo	referente a processos de trabalho discreto e sistemas operacionais
	Funcional	aplicação do processo que compara determinadas funções administrativas em duas ou mais organizações
	Desempenho	relativo às características quantificáveis em termos de preço, velocidade, confiabilidade
	Estratégica	avaliação envolvendo estratégica em vez de questões operacionais
Propósito de relacionamento	Competitiva	comparação para ganhar superioridade sobre os outros
	Colaborativa	comparação para o desenvolvimento de um ambiente de aprendizagem e compartilhamento do conhecimento

Fonte: Fong, Cheng e Ho (1998), adaptado.

A classificação proposta por Fong, Cheng e Ho (1998) tem relação direta com o objetivo do *benchmarking* seja no aspecto de qual será a referência, o que será comparado e qual o propósito da análise.

Em Anand e Kodali (2008), com base em Derors, Yusof e Salleh (2006), as aplicações de *benchmarking* são classificadas conforme a origem:

- **acadêmico** - modelos desenvolvidas por acadêmicos e pesquisadores, por meio de suas pesquisas, conhecimento e experiências. Nestes modelos, o pesquisador tem um foco teórico e conceitual, que pode ou não ter sido implementado e validado com uma aplicação prática real;
- **consultor** - modelos desenvolvidos com base na experiência na prestação de consultoria para organizações em projetos de benchmarking. Estes modelos, geralmente, são testados e validados com a implementação na organização do cliente e, portanto, a abordagem tende a ser orientada à prática;
- **organização** - modelos desenvolvidos pelas próprias organizações com base nos processos, experiência e conhecimento interno. Estes tendem a ser altamente heterogêneos, uma vez que cada organização em vários aspectos (foco do negócio, mercado, produtos, processos).

Os pesquisadores Anand e Kodali (2008, p. 287) classificaram as 35 publicações clássicas sobre modelos de *benchmarking* até a década de 2000 e chegaram aos seguintes números: 17 (48%) são do tipo Consultor, 9 (26%) tipo Acadêmico e 9 (26%) tipo Organização. A maioria dos modelos, na pesquisa, foi desenvolvido por consultores, mostrando o predomínio dos modelos de ordem prática e aplicados nas empresas.

2.4 AS APLICAÇÕES DE BENCHMARKING

Com relação as técnicas analíticas utilizadas em *benchmarking*, segundo Moriarty (2011), destacam-se a DEA (*Data Envelopment Analysis*), a AHP (*Analytical Hierarchy Process*), a PCA (*Principal Component Analysis*), a CFA (*Common Factor Analysis*). Neste mesmo estudo, Moriarty desenvolveu um *framework* teórico, que estabelece as condições necessárias para que o processo de *benchmarking* seja efetivo nas empresas, por intermédio de axiomas e condições lógicas.

De certa forma, para validar a pesquisa de Moriarty (2011), os critérios para a busca sistematizada dos trabalhos (descritos na seção 4.2 PROCEDIMENTOS METODOLÓGICOS) foi identificar trabalhos correlacionados com a proposição desta tese, que são publicações com aplicação prática de *benchmarking*, utilizando-se de uma técnica quantitativa, independente da área de conhecimento. Buscou-se, assim, trabalhos com duas características básicas: uso de técnicas quantitativas e o desenvolvimento de estudos de casos.

Dos 176 artigos pré-selecionados de forma sistematizada, somente 17 (9,7%) estão diretamente correlacionados com os propósitos de utilização da TRI como técnica quantitativa de *benchmarking*. Entretanto, esses utilizam outras técnicas quantitativas. No Quadro 3 estão as principais informações destes dezessete artigos, contendo a técnica utilizada, autor, título do trabalho e ano de publicação.

Quadro 3 - Artigos selecionados de aplicações de *benchmarking*.

Técnica	Autor	Título	Ano
Fuzzy AHP, TOPSIS	Gülçin BÜYÜKÖZKAN	A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry	2012
Fuzzy AHP	Gülçin BÜYÜKÖZKAN	Strategic analysis of healthcare service quality using fuzzy AHP methodology	2011
DEA	Nazim N. HABIBOV	Comparing and contrasting poverty reduction performance of social welfare programs across jurisdictions in Canada using Data Envelopment Analysis (DEA): An exploratory study of the era of devolution	2010
MCDA	Fotios PASIOURAS	Multicriteria classification models for the identification of targets and acquirers in the Asian banking sector	2010
ANOVA, HLM	Michael NAOR	The globalization of operations in Eastern and Western countries: Unpacking the relationship between national and organizational culture and its	2010

		impact on manufacturing performance	
IMMF	Devendra D. POTNIS	Measuring e-Governance as an innovation in the public sector	2010
GP, TOPSIS	Mohammad Reza MEHREGAN	An optimisational model of benchmarking	2010
DEA, TFP	Maria Luisa CORTON	Benchmarking Central American water utilities	2009
QFD, AHP	A. ANDRONIKIDIS	The application of quality function deployment in service quality management	2009
DEA	Hui Boon TAN	Relative efficiency measures for the knowledge economies in the Asia Pacific region	2008
DEA	Preethy NAYAR	Data Envelopment Analysis Comparison of Hospital Efficiency and Quality	2008
DEA	Amer HAMDAN	Evaluating the efficiency of 3PL logistics operations	2008
SEM	S.Y. SOHN	Structural equation model for the evaluation of national funding on R&D project of SMEs in consideration with MBNQA criteria	2007
MCDM	William HO	Multiple criteria decision-making techniques in higher education	2006
CI	Hariolf GRUPP	Indicators for national science and technology policy: how robust are composite indicators?	2004
DEA	Wai Ho LEONG	Measuring the technical efficiency of banks in Singapore for the period 1993-99: an application and extension of the Bauer et al. (1997) technique	2003
DEA	Mike DONNELLY	A radical scoring system for the European Foundation for Quality Management Business Excellence Model	2000

Fonte: Elaborado pelo Autor (2011).

Com base no Quadro 3, identificou-se que a técnica mais utilizada foi a DEA (7; 41,2%), (HABIBOV, 2010; CORTON, 2009; TAN, 2008; NAYAR, 2008; HAMDAN, 2008; LEONG, 2003; DONNELLY, 2000), seguida da AHP (3; 17,6%) (BÜYÜKÖZKAN, 2012; BÜYÜKÖZKAN, 2011; ANDRONIKIDIS, 2009). As demais técnicas utilizadas apareceram de forma individualizada: ANOVA (*Analysis of Variance*) e HLM (*Hierarchical Linear Modeling*) (NAOR, 2010); GP (*Goal Programming*) e TOPSIS (*Technique for Order Preference by Similarity to Ideal Solution*) (MEHREGAN, 2010); IMMF (*Innovation Management Measurement Framework*) (POTNIS, 2010); MCDA (*Multicriteria Decision Aid*) (PASIOURAS, 2010); SEM (*Structural Equation Model*) (SOHN, 2007); MCDM (*Multiple-Criteria Decision-Making*) (HO, 2006); CI (*Composite Indicators*) (GRUPP, 2004).

Chega-se, então, que as duas técnicas (DEA e AHP) são aplicadas em quase 60% dos artigos selecionados, corroborando com os resultados de Moriarty (2011) e, em sendo assim, estas serão o foco comparativo com a TRI, como as técnicas quantitativas mais aderentes aos propósitos deste trabalho. Essa discussão está apresentada nas seções a seguir.

2.4.1 DEA

A técnica DEA, conhecida como análise por envoltória de dados, é uma abordagem multivariada, contemporânea ao processo de *benchmarking* (CHARNES; COOPER; RHODES, 1978), orientada a dados, que tem como principal objetivo avaliar o desempenho de um conjunto de entidades pares, denominadas de *Decision Making Units* (DMU). As DMU, objetivamente, convertem múltiplas entradas em várias saídas. São exemplos de DMU: universidades, bancos, cidades, estados, países, indústrias, hospitais, empresas em geral. Uma grande variedade de aplicações da DEA pode ser realizada na avaliação dos desempenhos de diversos tipos de DMU, em muitos contextos e locais diferentes (COOPER; SEIFORD; ZHUS., 2004).

Na sua essência, a DEA realiza uma combinação de todas as variáveis de entrada e de saída das DMU em um única medida de eficiência em um intervalo numérico de $[0; 1]$, onde “0” representa a total ineficiência e “1” a total eficiência. (LEONG, 2003). Matematicamente, a expressão que traduz essa essência está representada em (2.1):

$$Eficiência = \frac{\sum_s \text{saídas ponderadas}}{\sum_m \text{entradas ponderadas}} \quad (2.1)$$

Considera-se que existem ao todo de n DMU com um total de m entradas e s saídas, sendo que modelagem matemáticas para a proposta da expressão (2.1) para a determinação da eficiência foi desenvolvida por Charnes, Cooper e Rhodes (1978) e está apresentada em (2.2):

$$\max \frac{\sum_{k=1}^s v_k y_{kp}}{\sum_{j=1}^m u_j x_{jp}}, \text{ onde } \frac{\sum_{k=1}^s v_k y_{kp}}{\sum_{j=1}^m u_j x_{jp}} \leq 1 \text{ e } v_k, u_j \geq 0 \quad (2.2)$$

onde:

- v_k são os pesos das k saídas;
- u_j são os pesos das j entradas;
- y_{kp} é o conjunto de variáveis das k saídas das p DMU;
- x_{jp} é o conjunto de variáveis das j entradas das p DMU;
- p é o índice para as DMU;
- k é o índice para as variáveis de entradas;
- j é o índice para as variáveis de saídas;

Para a resolução da expressão (2.2), faz-se necessário a transformação em um problema de programação fracional linear, conforme apresentado na expressão (2.3):

$$\max e_k = \sum_{k=1}^s v_k y_{kp} \quad (2.3)$$

com as seguintes restrições:

- $v_k, u_j \geq 0$
- $\sum_{j=1}^m u_j x_{jp} = 1$
- $\sum_{k=1}^s v_k y_{kp} - \sum_{j=1}^m u_j x_{jp} \leq 0$
- $\sum_{k=1}^s v_k = \sum_{j=1}^m u_j$

A principal vantagem da DEA é a sua orientação empírica e a ausência da necessidade de hipóteses *a priori*, principalmente com relação às abordagens estatísticas, tais como as formas tradicionais de análise de regressão estatística, resultou na sua utilização em uma grande gama de estudos envolvendo análises de eficiência nos setores governamentais e sem fins lucrativos, nos setores que possuem regulações e, também, no setor privado (COOPER; SEIFORD; TONE,

2000). Assim, entende-se que, a DEA é uma metodologia direcionada para as fronteiras dos dados, em vez das tendências centrais, como nas regressões estatísticas. Com esta perspectiva, o DEA demonstra conseguir descobrir relações implícitas que não podem ser explicitadas por outras metodologias, pois o objetivo é dizer que uma DMU é mais eficiente do que outra DMU. Isto é obtido sem a exigência de pressupostos explicitamente formuladas, como em modelos de regressão linear e não linear (COOPER; SEIFORD; TONE, 2000).

Na aplicação de Habibov (2010), na análise da desempenho da redução da pobreza por meio de programas em províncias canadenses, a DEA fornece um estudo exploratório, demonstrando que a técnica pode fornecer conhecimentos teóricos e práticos. Mas uma limitação importante na análise é que as causas precisas das diferenças entre os programas não podem ser explicitados pela técnica, bem como as causas das alterações na eficiência ao longo do tempo.

Em Corton (2009), está apresentada uma análise da eficiência das empresas fornecedoras de águas em seis países na América Central e este ressalta que existem três fatores principais para que os resultados da técnica sejam confiáveis: a homogeneidade do conjunto das DMU, a escolha dos conjuntos de variáveis de entrada e saída utilizadas na comparação e a qualidade dos dados destas variáveis.

No trabalho de Nayar (2008), a DEA apresenta-se como uma ferramenta de *benchmarking* adequada para medir a qualidade em hospitais. Uma dificuldade foi a falta de dados em alguns casos, que obrigatoriamente promoveu a eliminação de algumas variáveis de entrada na análise. Em Tan (2008), a aplicação foi para mensurar a eficiência de países em transformar o conhecimento em desenvolvimento econômico e o resultado demonstra que quatro países emergentes são relativamente ineficientes, frente a países desenvolvidos. Segundo o autor, mais importante do que o resultado são as possibilidades de análise com foco no desenvolvimento econômico de forma ampla, não somente baseada em indicadores financeiros e de infraestrutura física. Em Hamdan (2008), na análise de operações logísticas de depósitos de uma companhia, o autor insere restrições sintonizadas aos objetivos dessa empresa, definidas por especialistas internos e obtém um resultado considerado mais adequado com relação ao uso do modelo de Cooper, Seiford e Tone (2000) sem essas restrições. Sendo assim, percebe-se a dependência dos resultados da análise da eficiência em função de restrições adequadas ao problema, definidas por especialistas.

Nos dois últimos trabalhos selecionados com aplicações da DEA têm como objetivo comparar modelos de desempenho elaborados para avaliar a eficiência do sistema bancário de Singapura no período de 1993-1999 (LEONG, 2003) e comparar o ranking gerado pelo escore clássico do sistema de avaliação de gestão da qualidade usado no EFQM-BEM (*European Foundation for Quality Management's Business Excellence Model*) (DONNELLY, 2000). Em ambos, a DEA apresentou melhores resultados.

Com base nos resultados das aplicações da DEA, conclui-se que a facilidade de aplicação prática, que por ser uma técnica não paramétrica, não necessita de premissas de distribuição nos dados nas variáveis de entrada e saída. O desempenho é medido em função da relação entre as variáveis de entrada e saída, sendo que a técnica não privilegia as DMUs com mais recursos disponíveis, isto significa que os resultados obtidos na análise de desempenho são imunes à quantidade de recursos. Essas são as principais vantagens desta técnica. As limitações identificadas foram que o resultado dos desempenhos dependem da definição das restrições utilizados para resolução da programação linear, da necessidade da escolha adequada dos conjuntos de variáveis de entrada e de saída, e principalmente, a não possibilidade de interpretação dos resultados obtidos em função das variáveis de entrada e saída escolhidas.

2.4.2 AHP

A técnica AHP é uma eficaz abordagem quantitativa no desenvolvimento de *benchmarking* como apoio na aplicação da *Quality Function Deployment* (QFD), uma vez que possibilita que decisões complexas não se baseiem exclusivamente no instinto dos especialistas. Como uma das principais necessidades no processo do *benchmarking* é a definição da priorização das características a serem implementadas, a AHP permite essa priorização de forma quantitativa (QFD INSTITUTE, 2012).

A QFD é um sistema voltado para a qualidade e foi desenvolvido para reduzir a distância crescente entre os desenvolvedores e usuários finais, promovendo uma interface pessoal de produção moderna e de negócios. O foco é atender as verdadeiras necessidades do cliente (usuário final) com design, desenvolvimento, engenharia, fabricação e serviços de funções, sendo um sistema de qualidade que implementa elementos do pensamento sistêmico com elementos de Psicologia e Epistemologia (conhecimento) (QFD INSTITUTE, 2012).

Em Carnevalli e Miguel (2008) está uma ampla revisão de trabalhos recentes (2002 a 2006) em QFD, classificação, escopo, definições e benefícios, pré-requisitos usados, principais dificuldades no uso, metodologias encontradas e recomendações para redução dessas características. Segundo esses autores, QFD “é um importante método de desenvolvimento de produto, dedicada a traduzir os requisitos do cliente em atividades de desenvolvimento de produtos e serviços” (p. 737).

Para Büyüközkan (2011) e (2012), o uso da avaliação da qualidade do serviço no setor da saúde com AHP e com a inserção da lógica difusa (*fussy*), maximiza a assertividade de técnica, minimizando a incerteza humana na atribuição dos pesos necessárias. Em Andronikidis (2009), o uso do AHP em conjunto com o QFD foi para tornar a análise mais precisa, na atribuição dos pesos na *House of Quality* (HOQ). Em ambos os trabalhos, a AHP foi uma técnica utilizada para selecionar as iniciativas a serem buscadas, ajudando os gestores e desenvolvedores para que, ao combinarem todas essas informações, possam tomar decisões mais apropriadas. A técnica permite reduzir decisões complexas a uma série de comparações por pares, e em seguida, sintetizar os resultados em uma matriz de priorização, conforme apresentada na planilha disponível em ASQ (2013).

A principal vantagem da AHP é, portanto, possibilitar a priorização das melhores práticas de *benchmarking* com a minimização da dependência de especialistas (juízes) para a priorização, mas não elimina totalmente a dependência da subjetividade desses juízes. E essa dependência da subjetividade, em que consiste a principal limitação da técnica.

3 - TEORIA DA RESPOSTA AO ITEM

Neste capítulo, o objetivo é apresentar a origem e as definições pertinentes ao tema, bem como os principais modelos estatísticos da TRI. A apresentação das aplicações da TRI classificadas por área de conhecimento fazem o fechamento deste capítulo. No início deste é explorada a teoria da medida.

3.1 TEORIA DA MEDIDA

A teoria da medida discute como os fenômenos naturais podem ser representados por números, e segundo Pasquali (2003), são três os problemas que esta teoria tem que resolver:

- problema da representação - garantia da legitimidade de representar, por meio de números, as propriedades dos fenômenos naturais;
- problema da unicidade da representação - garantia que o número é a única/melhor forma de representação das propriedades dos fenômenos naturais;
- problema do erro - estes causados pelos instrumentos de medição, pelos observadores e os aleatórios.

Segundo Pasquali (2003, p.26), são duas perguntas que devem ser respondidas quando se quer utilizar de números para quantificar os fenômenos naturais: “1 - *É legítimo utilizar o número para descrever os fenômenos da ciência?* 2 - *É útil, vantajoso, utilizar o número para descrever os fenômenos da ciência?*”.

Intrinsecamente associado ao número, que representa um fenômeno natural, tem-se a escala de medida, e dependendo do que se quer medir, deve-se utilizar a escala mais apropriada. A construção da escala possibilita a mensuração dos fenômenos naturais, ou seja, que as características destes possam ser transformadas em números. Alguns fenômenos naturais são facilmente medidos diretamente, como peso, altura e idade. Entretanto, outros como inteligência, qualidade de vida e depressão não podem ser medidos diretamente. Esses últimos são chamados de traços latentes e para medi-los, é necessário que se construa uma escala psicométrica (PASQUALI, 1998).

O procedimento para criação teórica de uma escala, para mensurar um traço latente, é descrito em passos em Pasquali (1998):

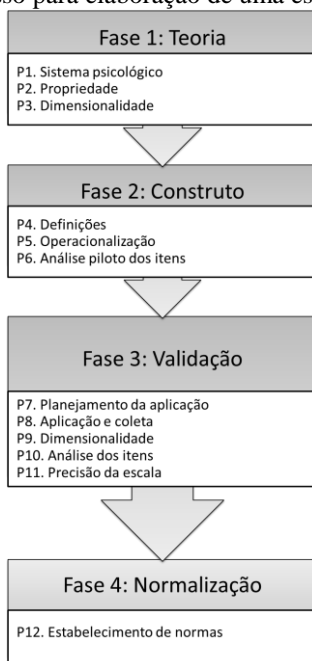
- **1º sistema psicológico** - levantamento de materiais teóricos sobre o traço latente;
- **2º propriedades** - levantamento dos atributos com base nas teorias do traço latente;
- **3º análise da dimensionalidade** - identificação de fatores a investigar com base na teoria levantada;
- **4º definição do teste** - pode ser realizada em duas formas: constitutiva, com base em outros testes identificados no levantamento teórico; operacional, com definições que permitem medir o construto;
- **5º operacionalização do teste** - construção dos itens levando em consideração as suas fontes, oriundas da literatura, de entrevista com especialistas; regras de construção, definição de critérios; quantidade de itens;
- **6º análise teórica dos itens** - realizada em duas etapas: análise semântica, compreensão dos itens; análise dos juízes operacional, que possibilita a associação dos itens ao traço latente.

Os seis passos descritos permitem elaborar o teste de forma teórica, e na sequência, tem-se as fases de validação e normatização. Essas duas últimas, são de aplicação do teste, que consistem do uso da TRI como forma de operacionalização da medição do traço latente, que será explorado nos demais itens deste capítulo. O processo completo de construção de uma escala psicométrica está apresentado na Figura 6.

3.2 ORIGEM E CONCEITOS DA TRI

A TRI é uma técnica quantitativa que permite a medição de traços latentes, como por exemplo, o conhecimento de matemática e português, o nível de *stress* e depressão, e, também, a competitividade de produtos de empresas, foco deste trabalho. A TRI é um processo apoiado por um conjunto de modelos estatísticos, inicialmente desenvolvidos e aplicados somente nas áreas da psicologia e da educação. A ideia moderna da TRI foi apresentada por Thurstone em 1925, num artigo cujo título é: *A Method of Scaling Psychological and Educational Tests* (BOCK, 1997). A Figura 7 representa o estudo do Thurstone, em que o mesmo construiu um gráfico dos percentuais de acerto dos alunos em função das idades desses.

Figura 6 - Processo para elaboração de uma escala psicométrica.

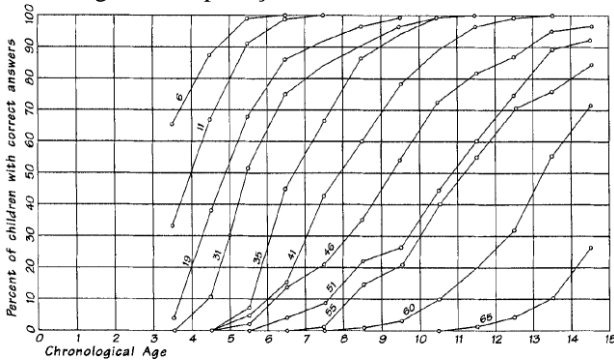


Fonte: Pasquali (1998), adaptado.

Na sua essência, a TRI é uma técnica quantitativa que possui modelos estatísticos com o objetivo de medir o traço latente, por intermédio da construção de uma escala de medição, que de forma indireta, quantifica esse traço. A construção da escala é possível por intermédio da definição de itens adequados para a medição do respectivo traço latente. Finalizada a análise, o processo é realimentado, com a possibilidade de revisão dos itens e melhoria do teste, aumentando a importância da tradução do traço latente em itens válidos para a medição deste (PASQUALI, 2003). Esquemáticamente, o princípio da TRI está apresentado na Figura 8.

O que seria um traço latente? Como definir os itens a utilizar? Como se constrói o teste e a escala? A resposta a estas três perguntas constituem os conceitos necessários para uma aplicação consistente da TRI (ANDRADE; TAVARES; VALLE, 2000). Com base na Figura 8, percebe-se a importância da tradução do traço latente em itens válidos para a medição deste.

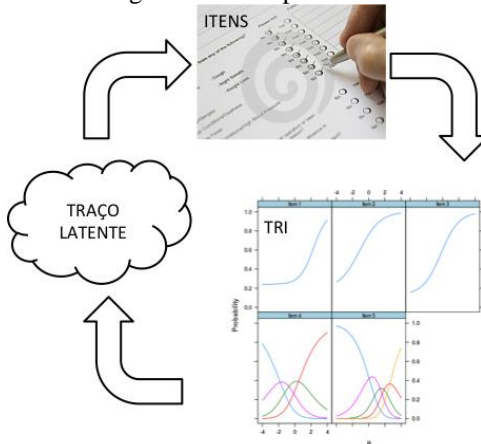
Figura 7 - Aplicação de Thurstone em 1925.



Fonte: Bock (1997).

O traço latente pode ser conceituado como uma característica oculta que não pode ser medida diretamente, mas que é possível ser observada nas respostas de um teste (PASQUALI, 2003). As duas formas básicas de modelagem do traço latente, segundo Guttman (1950), é a forma cumulativa e, que conforme Coombs (1964), é a forma de desdobramento (*unfolding*).

Figura 8 - Princípio da TRI.



Fonte: Elaborado pelo Autor (2011).

A forma cumulativa relaciona-se, principalmente, a medições de habilidades, como por exemplo o conhecimento de um assunto, pois existe uma relação direta entre o número de acertos dos itens em uma

prova com o conhecimento do indivíduo, sendo esse aspecto que define o princípio de acumulação dos itens que modelam o traço latente. Essa forma também aplica-se a medições psicológicas, como depressão e qualidade de vida, pois quanto maior o número de itens confirmados com relação ao respectivo diagnóstico, maior o nível da depressão ou da qualidade de vida.

Na forma de desdobramento, as aplicações focam-se, principalmente, na medição de atitudes, em que a resposta de um item relaciona-se a uma reação, comportamento ou opinião do indivíduo com relação a certos objetos, instituições, conceitos, pessoas ou situações, representados por itens. Nesse caso, o traço latente representa se o indivíduo posiciona-se favorável ou não a esses itens.

Para a medição do traço latente é necessário a elaboração de itens que representem o traço latente cumulativo ou não cumulativo (desdobramento), que se deseja medir. O traço latente independe dos itens e não deve alterar quando esses são aprimorados, desde que os novos itens, consigam da mesma forma medir o traço latente em questão.

3.3 OS MODELOS DA TRI

Para a definição do modelo estatístico da TRI mais adequado para modelar um traço latente, tem-se que levar em consideração quatro aspectos: (1) a natureza do item; (2) o número de populações envolvidas; (3) a quantidade de traços latentes (ANDRADE; TAVARES; VALLE, 2000); (4) e a característica cumulativa ou não cumulativa do traço latente (BORTOLOTTI, 2003).

A definição se o traço latente representa, por exemplo um conhecimento (cumulativo) ou uma atitude (desdobramento) é o início do processo. Segue-se, então, para o número de dimensões do traço latente, onde este aspecto é que define a dimensionalidade do modelo, se o mesmo modelará um traço unidimensional ou multidimensional.

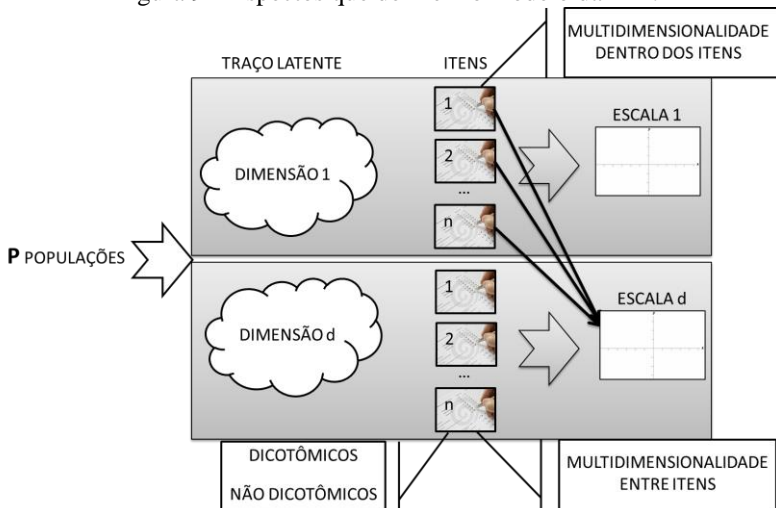
Chega-se na definição dos itens, em que deve existir uma sintonia entre os itens elaborados e o traço latente, uma vez que os itens têm que representar o traço latente nos dois aspectos já mencionados. Pode-se dividir o tipo dos itens em dois grupos: dicotômicos e politômicos. Os itens dicotômicos são também conhecidos como binários e representam somente duas possibilidades, como por exemplo: certo ou errado, presente ou ausente, concorda ou discorda. Já os itens politômicos, tem que ter no mínimo três opções como possibilidades, sendo que essas

podem ser nominais ou ordinais. A diferença entre estes dois (nominais e ordinais) tipos de itens é que os ordinais possuem uma relação de ordem entre as categorias. A escala *likert* (LIKERT, 1932) é o exemplo clássico de item do tipo ordinal. Já os nominais, representam categorias sem uma ordem pré-estabelecida, como nas respostas de itens com múltiplas escolhas, em que somente uma está correta. Podem existir também itens quantitativos, cuja a resposta ao mesmo é um número, estes itens podem ser transformados nos tipos binários, ordinal ou nominal.

Mesmo com todas as possibilidades de tipos de itens, estes têm que ser construídos de acordo com o traço latente que se deseja medir, sintonizados à dimensionalidade do modelo. Em caso de modelos multidimensionais, pode ocorrer um item estar relacionado a uma ou mais dimensões.

Com as definições já feitas, entende-se que o construto está pronto, sabe-se se é cumulativo ou desdobramento, se unidimensional ou multidimensional, os tipos e quantidades de itens. Com o construto elaborado, pode-se aplicar o instrumento de coleta a uma ou mais populações, quando se tem a necessidade de comparação de grupos distintos. O diagrama da Figura 9 representa os aspectos que definem o modelo da TRI a ser utilizado.

Figura 9 - Aspectos que definem o modelo da TRI.



Fonte: Elaborado pelo Autor (2011).

Com base na Figura 9, um traço latente pode ser representado por d dimensões, com seus respectivos n itens e respectivas escalas de medição. Pode ocorrer, que itens são construídos para a medição em duas dimensões simultaneamente de um traço latente, sendo que esta situação é denominada multidimensionalidade dentro dos itens. Quando os itens medem as dimensões de um traço latente isoladamente, tem-se a multidimensionalidade entre itens (BABCOCK, 2009).

A dimensionalidade do traço latente agrupa os modelos estatísticos da TRI em dois grupos: modelos unidimensionais e multidimensionais. Com relação ao tipo dos itens, também surgem dois grupos: modelos com itens dicotômicos e com itens politômicos.

Com o objetivo de exemplificar a aplicação dos aspectos apresentados nos modelos da TRI, aqui são apresentados somente os principais. Os modelos logísticos unidimensionais com itens dicotômicos são os mais simples e, também, os mais utilizados, possuindo um, dois ou três parâmetros, denominados ML1, ML2 e ML3, respectivamente. Estes modelos estão apresentados nas expressões (3.1), (3.2) e (3.3) (ANDRADE; TAVARES; VALLE, 2000):

$$\text{ML3: } P(U_{ij} = 1/\theta_j) = c_i + \frac{(1-c_i)}{1+e^{-a_i(\theta_j-b_i)}} \quad (3.1)$$

$$\text{ML2: } P(U_{ij} = 1/\theta_j) = \frac{1}{1+e^{-a_i(\theta_j-b_i)}} \quad (3.2)$$

$$\text{ML1: } P(U_{ij} = 1/\theta_j) = \frac{1}{1+e^{-(\theta_j-b_i)}} \quad (3.3)$$

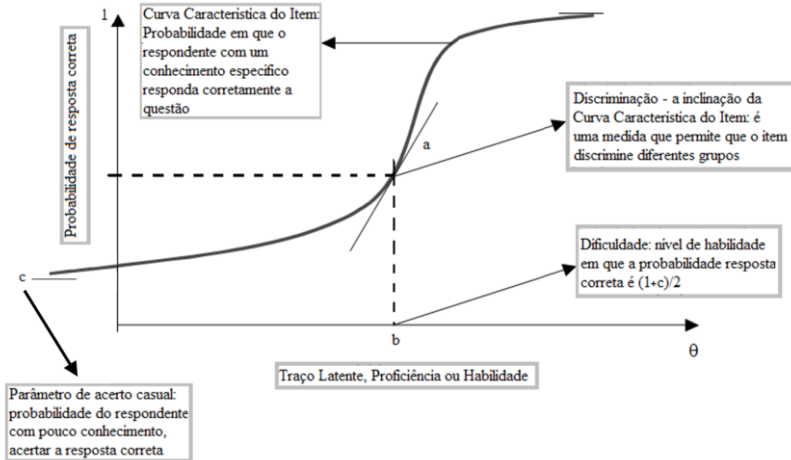
onde:

- θ_j é a proficiência (traço latente) do indivíduo j ;
- $P(U_{ij} = 1 / \theta_j)$ é a probabilidade de uma resposta correta do item i dada a proficiência θ_j do indivíduo j .
- a_i é o parâmetro de discriminação do item i , sendo $a_i > 0$;
- b_i é o parâmetro de posição do item i , sendo conhecido como parâmetro de dificuldade e representado na mesma escala do traço latente;
- c_i é o parâmetro da assíntota inferior do item i , que representa a probabilidade de indivíduo com baixa proficiência apresentar a resposta correta (acerto casual).

Os modelos ML2 e ML1 são casos especiais do ML3. No ML2, o parâmetro da assíntota é nulo ($c=0$). No ML1, além do parâmetro da

assíntota nulo, o parâmetro de discriminação é o mesmo para todos os itens. Na Figura 10, está apresentada a curva característica de um item de um modelo ML3, onde cada parâmetro do modelo está identificado.

Figura 10 - Curva Característica de um Item.



Fonte: Andrade, Tavares e Valle (2000), adaptado.

A generalização dos modelos logísticos em (3.1), (3.2) e (3.3) para duas ou mais populações foi desenvolvido por Bock e Zimowski (1997) é apresentado em (3.4):

$$P(U_{ijk} = 1/\theta_{jk}) = c_i + \frac{(1-c_i)}{1+e^{-a_i(\theta_{jk}-b_i)}} \quad (3.4)$$

onde:

- k é o índice de identificação das populações;
- $P(U_{ijk} = 1 / \theta_{jk})$ é a probabilidade de uma resposta correta do item i dada a proficiência θ_j do indivíduo j , da população k ;
- os demais parâmetros (a , b e c) possuem as mesmas interpretações de (1) generalizados para k populações.

O modelo unidimensional com itens politômicos mais utilizado na TRI é o Modelo de Resposta Gradual (MRG - *Graded Response Model*) (SAMEJIMA, 1969). Esse modelo tem o intuito de obter mais informação dos respondentes do que simplesmente se eles deram respostas corretas ou incorretas, ou se concordaram ou não. Na

expressão (3.5) está o modelo MRG:

$$\text{MRG: } P_{i,k}^+(\theta_j) = \frac{1}{1+e^{-a_i(\theta_j-b_{i,k})}} \quad (3.5)$$

onde:

- k é o índice de identificação das opções.
- $P_{i,k}^+(\theta_j)$ é a probabilidade de um indivíduo j escolher uma particular categoria k ou outra de maior ordem do item i ;
- os demais parâmetros (a e b) possuem as mesmas interpretações de (3.1) generalizados para k populações.

Outros modelos particulares decorrentes do Samejima são apresentados a seguir: o Modelo de Escala Gradual (MEG - *Rating Scale Model*) (ANDRICH, 1978), o Modelo de Crédito Parcial (MCP - *Partial Credit Model*) (MASTERS, 1982) e o Modelo de Crédito Parcial Generalizado (MCPG - *Generalized Partial Credit Model*) (MURAKI, 1992). Estes modelos estão apresentados em (3.6), (3.7) e (3.8), respectivamente:

$$\text{MEG: } P_{i,k}(\theta_j) = \frac{1}{1+e^{-a_i(\theta_j-b_i+d_k)}} - \frac{1}{1+e^{-a_i(\theta_j-b_i+d_{k+1})}} \quad (3.6)$$

$$\text{MCP: } P_{i,k}(\theta_j) = \frac{e^{-\sum_{u=0}^k a_i(\theta_j-b_{i,u})}}{\sum_{u=0}^{m_i} e^{-\sum_{v=0}^u a_i(\theta_j-b_{i,v})}} \quad (3.7)$$

$$\text{MCPG: } P_{i,k}(\theta_j) = \frac{e^{-\sum_{u=0}^k a_i(\theta_j-b_{i,u})}}{\sum_{u=0}^{m_i} e^{-\sum_{v=0}^u a_i(\theta_j-b_{i,v})}} \quad (3.8)$$

O MEG é adequado para itens com categorias de resposta ordenadas, porém é feita a suposição que os escores das k categorias de todos os itens são igualmente espaçados por d_k , o parâmetro de categorias. O MCP é um modelo para análise de respostas obtidas de duas ou mais categorias ordenadas, pertencente à família de modelos de *Rasch*, que possuem o mesmo parâmetro de discriminação para todos os itens. O MCPG é igual ao MCP, porém com parâmetros de discriminação a diferentes para os itens i .

Os modelos apresentados são exemplos de modelos acumulativos da TRI e a utilização dos mesmos dependerá da modelagem da competitividade do produto, com base nas suas características. Poderão

existir produtos com traços latentes unidimensionais ou multidimensionais, com características dicotômicas e/ou politômicas.

3.4 O PROCESSO DE APLICAÇÃO DA TRI

Dando prosseguimento ao processo de construção de escalas psicométricas descrito em 3.1 (Figura 6) e anterior à aplicação da TRI, tem-se inicialmente a validação do construto por meio de um teste piloto, seguindo com o planejamento da aplicação e a coleta dos dados propriamente dita. Estas etapas são pertinentes à aplicação de instrumentos de coleta para quaisquer fins e não serão explorados neste trabalho.

A análise da TRI tem como objetivo criar uma escala para medir o traço latente, que não pode ser medido de forma direta, e a mesma inicia após a etapa de criação e validação do instrumento de medida, que representa o traço latente, o construto. A análise é subdividida em duas etapas: estimação (calibração) dos parâmetros do modelo estatístico, como os apresentados na seção 3.3, e a estimação da proficiência, a quantificação do traço latente. Por facilidades computacionais, os processos de estimação são realizados na escala (0,1), na qual “0” representa a média do traço latente dos respondentes utilizados na criação da escala e “1” a medida de desvio padrão.

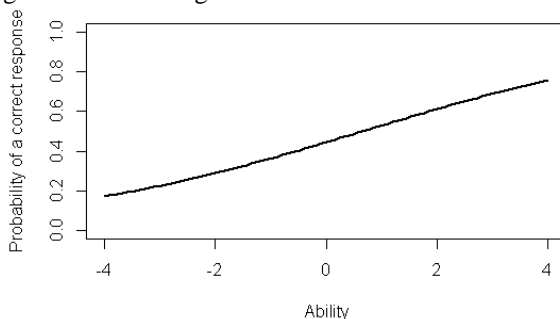
3.4.1 Calibração dos parâmetros dos itens

Os parâmetros dos itens independem dos indivíduos respondentes e a proficiência independe dos itens apresentados, não devendo mudar com a troca por outros itens válidos, que meçam o mesmo traço latente. Idealmente, o instrumento deve possuir itens com elevados graus de discriminação (parâmetro a) e com diferentes parâmetros de dificuldade (parâmetro b) em toda a escala (BAKER; KIM, 2004). O ajuste do modelo da TRI equivale à interpretação do α de Cronbach, que mede a consistência interna. Entretanto, o erro é melhor tratado nos modelos da TRI, pois os mesmos são individuais na estimação dos parâmetros para cada item do teste (ANDRADE; TAVARES; VALLE, 2000).

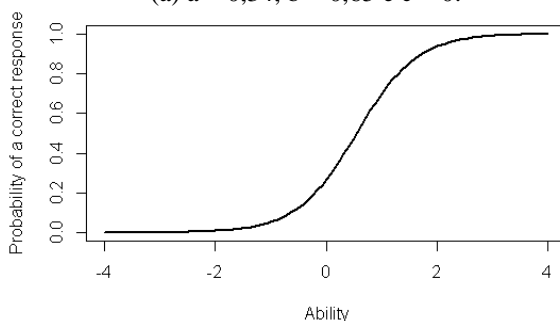
A confirmação da qualidade dos itens é realizada nesta análise, em que os critérios a serem utilizados são, tomando como referência a escala (0,1): parâmetro de discriminação (a) maior ou igual a 0,75 e parâmetro de dificuldade (b) estar no intervalo de [-3; +3]. Além da

análise dos parâmetros, é realizada a análise gráfica da Curva Característica do Item (CCI ou IRF - *Item Response Function*), na Figura 11.

Figura 11 - Análise gráfica de curvas características de itens.



(a) $a = 0,34$, $b = 0,65$ e $c = 0$.



(b) $a = 1,85$, $b = 0,55$ e $c = 0$.

Fonte: Elaborado pelo Autor (2012).

Os dois itens da Figura 11 têm praticamente o mesmo parâmetro de dificuldade, o gráfico (a) possui um parâmetro de discriminação abaixo do critério de qualidade. Já o gráfico (b) representa um item com qualidade. Caso o problema seja no parâmetro de discriminação, significa que o item não consegue discriminar grupos distintos e promover a construção da escala de proficiência, e se o problema for no parâmetro de dificuldade, a posição do item pode deixá-lo fora dessa escala.

Com relação ao parâmetro de dificuldade b vale uma ressalva, a escala varia de $-\infty$ a $+\infty$, e portanto qualquer valor de b é válido dentro dessa escala. Entretanto, na prática, se trabalha com a estimativa deste

parâmetro dentro de uma população e tem-se a expectativa que todos os valores de b estejam no intervalo $[-3; +3]$.

Em conjunto com as curvas características dos itens, utiliza-se de mais dois gráficos para análise dos itens, a Função de Informação do Item (FII ou IIF - *Item Information Function*) e a Função de Informação do Teste (FIT ou ITF - *Test Information Function*), cujas expressões para a construção das mesmas estão representadas em (3.9), (3.10) e (3.11), respectivamente (ANDRADE; TAVARES; VALLE, 2000):

$$I_i(\theta) = \frac{\left[\frac{d}{d\theta}P(U_{ij}=1/\theta)\right]^2}{P(U_{ij}=1/\theta)[1-P(U_{ij}=1/\theta)]} \quad (3.9)$$

$$I(\theta) = \sum_{i=1}^I I_i(\theta) \quad (3.10)$$

$$EP(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (3.11)$$

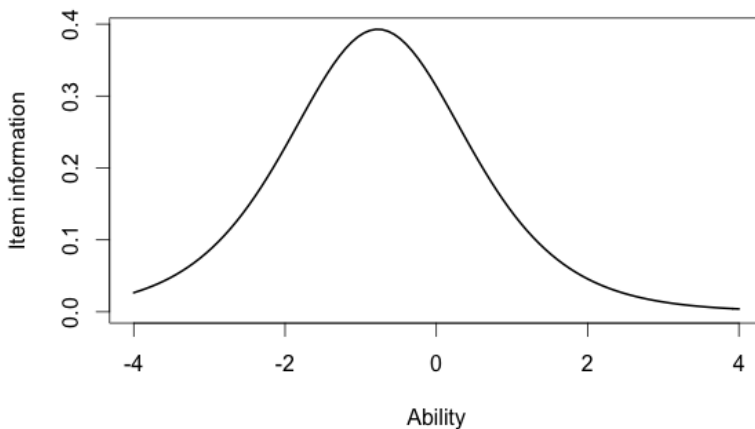
onde:

- θ é a proficiência do traço latente;
- $P(U_i = 1 / \theta)$ é a probabilidade de uma resposta correta do item i dada a proficiência θ ;
- $I_i(\theta)$ é a informação do item i no nível de proficiência θ ;
- $I(\theta)$ é a informação do teste no nível de proficiência θ ;
- $EP(\theta)$ é o erro padrão da estimação no nível de proficiência θ .

Na Figura 12 está um exemplo de uma FII, em que a maior contribuição do item está no entorno da proficiência igual a “-1”. Na Figura 13 está um exemplo de uma FIT, onde percebe-se uma concentração de informação no intervalo de proficiência de $[-2; 2]$.

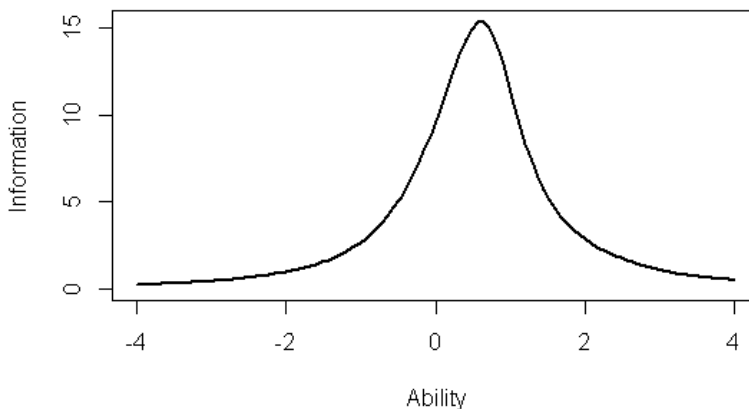
Com a finalização da calibração dos parâmetros dos itens, incluindo a verificação da qualidade dos itens, parte-se para a construção da escala. Isso é possível por meio da análise das estimativas dos parâmetros de dificuldade dos itens, que estão na mesma escala da proficiência. Essa análise é realizada mediante a definição de itens âncoras, que é uma das grandes vantagens da TRI, pois além de possibilitar a criação da escala, permite que a mesma seja interpretada com base nos itens (BEATON; ALLEN, 1992).

Figura 12 - Função de informação do item.



Fonte: Elaborado pelo Autor (2012).

Figura 13 - Função de informação do teste.



Fonte: Elaborado pelo Autor (2012).

Os critérios para definição de itens dicotômicos âncoras, com sucessivos valores de proficiências x e y , são dados em (3.12), (3.13) e (3.14):

$$P(U_{ij} = 1 / \theta_j = y) \geq 0,65 \quad (3.12)$$

$$P(U_{ij} = 1/\theta_j = x) \leq 0,50 \quad (3.13)$$

$$P(U_{ij} = 1/\theta_j = y) - P(U_{ij} = 1/\theta_j = x) \geq 0,30 \quad (3.14)$$

Pela expressão (3.12), a probabilidade de acerto, dada a proficiência y , tem que ser maior que 65%. Na expressão (3.13), a probabilidade de acerto, dada a proficiência x , tem que ser menor do que 50% e a diferença entre essas probabilidades (3.14) tem que ser maior do que 30%. Isto significa que deve existir a discriminação, formação de grupos, para se ter itens âncoras na escala.

3.4.2 Estimação do traço latente (proficiência)

A última etapa do processo de aplicação da TRI é a estimação da proficiência de cada indivíduo, o valor do traço latente de cada respondente, em que os principais métodos utilizados são Máxima Verossimilhança (MV) e os bayesianos com estimação pela moda *a posteriori* (MAP - *Maximum A Posteriori*) e pela média *a posteriori* (EAP - *Expected A Posteriori*) (ANDRADE; TAVARES; VALLE, 2000; PARTCHEV, 2012). Dentre estes, o mais utilizado é o EAP (MISLEVY; STOCKING, 1989; ANDRADE; TAVARES; VALLE, 2000) devido ao cálculo direto e por possuir o menor erro médio.

Se existir necessidade da comparação de grupos distintos, a TRI permite isso por meio da equalização. A condição necessária para realizar essa tarefa é que em ambos os grupos se esteja medindo o mesmo traço latente e que exista itens comuns nos dois construtos. Assim, a técnica permite colocar numa mesma escala, grupos de respondentes com níveis de proficiências distintos. A equalização pode ser realizada de duas formas: *a priori* e *a posteriori*. Na primeira, os coeficientes da equalização são determinados por um método desenvolvido para grupos múltiplos (ZIMOWSKI *et al.*, 2003). Na segunda, os coeficientes utilizam o princípio da simetria e são determinados pelo método *mean-sigma*, onde utiliza-se média (*mean*) e a variância (*sigma*) dos dois grupos a serem equalizados (ANDRADE; TAVARES; VALLE, 2000).

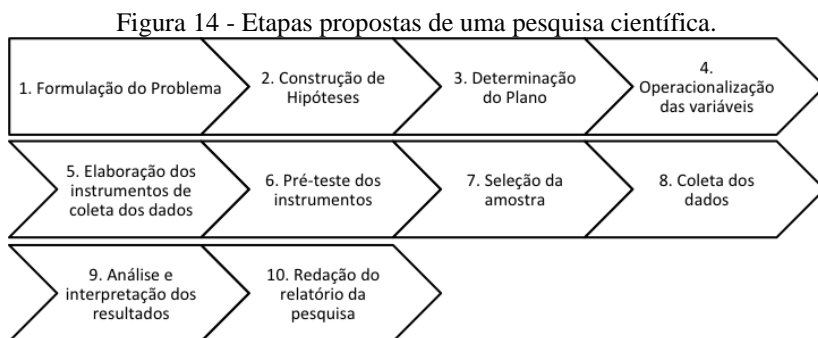
4 - METODOLOGIA

Este capítulo apresenta dois pontos da metodologia: a caracterização da pesquisa desenvolvida e os processos metodológicos para a elaboração do trabalho.

4.1 CARACTERIZAÇÃO DA PESQUISA

A ciência se faz quando o pesquisador aborda os fenômenos aplicando recursos técnicos, seguindo um método e apoiando-se em fundamentos epistemológicos (SEVERINO, 2007).

Para o desenvolvimento de trabalhos científicos, utiliza-se da pesquisa científica, que de forma organizada, permite provar ou não os objetivos e hipóteses inicialmente propostos pelo pesquisador. Para Gil (2002), “a pesquisa é desenvolvida mediante o concurso de conhecimentos disponíveis e a utilização cuidadosa de métodos, técnicas e outros procedimentos científicos”. A forma organizada é composta de etapas que, segundo uma proposição de Gil (2002), pode ser esquematizada em 10 etapas, conforme Figura 14.



Fonte: Gil (2002), adaptado.

Essas etapas não são absolutas e nem todas precisam necessariamente ser desenvolvidas, sendo que o pesquisador é quem deve decidir a melhor forma de adaptar à sua pesquisa (GIL, 2002). A customização das etapas tem que estar sintonizada ao problema de pesquisa.

Para atingir ao objetivo de pesquisa utiliza-se do “método”, mencionado no conceito inicial deste capítulo e corresponde à metodologia científica que, por meio de ferramentas apropriadas, permite o desenvolvimento da pesquisa científica com base em referenciais teóricos adequados.

Neste contexto, entende-se por metodologia científica: “Conjunto de procedimentos lógicos e de técnicas operacionais que permitem o acesso às relações causais constantes entre os fenômenos” (SEVERINO, 2007). Portanto, a metodologia científica é a ferramenta de trabalho do pesquisador que possibilita a este desenvolver a pesquisa científica. Somente com o conhecimento da pesquisa a ser desenvolvida é que se pode definir o procedimento metodológico.

A pesquisa pode ser classificada quanto à sua natureza, forma de abordagem, quanto aos seus objetivos e procedimentos técnicos (SILVA; MENEZES, 2005). No Quadro 4 estão as categorias para a classificação, segundo Silva e Menezes (2005):

Quadro 4 - Classificação de uma pesquisa científica.

Crítérios de classificação	Categorias	Descrição
Natureza	Básica	conhecimentos novos úteis sem aplicação prática prevista
	Aplicada	conhecimentos para aplicação prática e dirigidos à solução de problemas específicos
Forma de abordagem	Qualitativa	baseada na interpretação dos fenômenos e a atribuição de significados
	Quantitativa	quantização dos fenômenos com uso de recursos e de técnicas estatísticas
Objetivos	Exploratória	proporcionar maior familiaridade com o problema com vistas a torná-lo explícito ou a construir hipóteses
	Descritiva	descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis
	Explicativa	identificar os fatores que determinam ou contribuem para a ocorrência dos fenômenos
Procedimentos técnicos	Bibliográfico	elaborada a partir de material já publicado
	Documental	elaborada a partir de materiais que não receberam tratamento analítico
	Experimental	se determina um objeto de estudo, as variáveis que seriam capazes de influenciá-

		lo, as formas de controle e de observação dos efeitos que a variável produz no objeto
	Levantamento	interrogação direta das pessoas cujo comportamento se deseja conhecer
	Estudo de caso	estudo profundo e exaustivo de um ou poucos objetos de maneira que se permita o seu amplo e detalhado conhecimento
	<i>Expost-Facto</i>	quando o experimento se realiza depois dos fatos
	Ação	concebida e realizada em estreita associação com uma ação ou com a resolução de um problema coletivo
	Participante	se desenvolve a partir da interação entre pesquisadores e membros das situações investigadas

Fonte: Silva e Menezes (2005).

Com base nas possibilidades de classificação de uma pesquisa científica apresentadas no Quadro 4, percebe-se que o enquadramento mais apropriado para esta pesquisa é: natureza **aplicada**, forma de abordagem **quali-quantitativa**, objetivo **explicativo** e com os procedimentos técnicos **bibliográfica** e **experimental**.

4.2 PROCEDIMENTOS METODOLÓGICOS

A classificação da pesquisa científica, sugerida na seção anterior, resume como os procedimentos metodológicos serão aplicados durante o desenvolvimento da tese. A contribuição teórica do trabalho está sintonizada com a natureza e a abordagem, onde far-se-á uso da técnica quantitativa da TRI para a solução do problema específico de *benchmarking*. Já os objetivos e procedimentos técnicos resumem a aplicação prática do trabalho, que será o desenvolvimento bibliográfico e experimental, da criação da escala de competitividade de cursos do ensino superior.

Das etapas utilizadas no desenvolvimento da tese, com base na Figura 14, não foram contemplados os passos 5, 6 e 7, pois não serão realizados, pelo autor, a coleta de dados primários para a elaboração da aplicação prática. Com a adaptação dessa figura ao trabalho em desenvolvimento, chega-se a Figura 15.

Figura 15 - Etapas da pesquisa científica para este trabalho.



Fonte: Gil (2002), adaptado.

O desenvolvimento do referencial teórico da pesquisa, cuja sistematização se faz necessária, para promover a identificação de lacunas a serem investigadas e conseqüentemente novas contribuições para a área de conhecimento (VILLAS; MACEDO-SOARES; RUSSO, 2008)

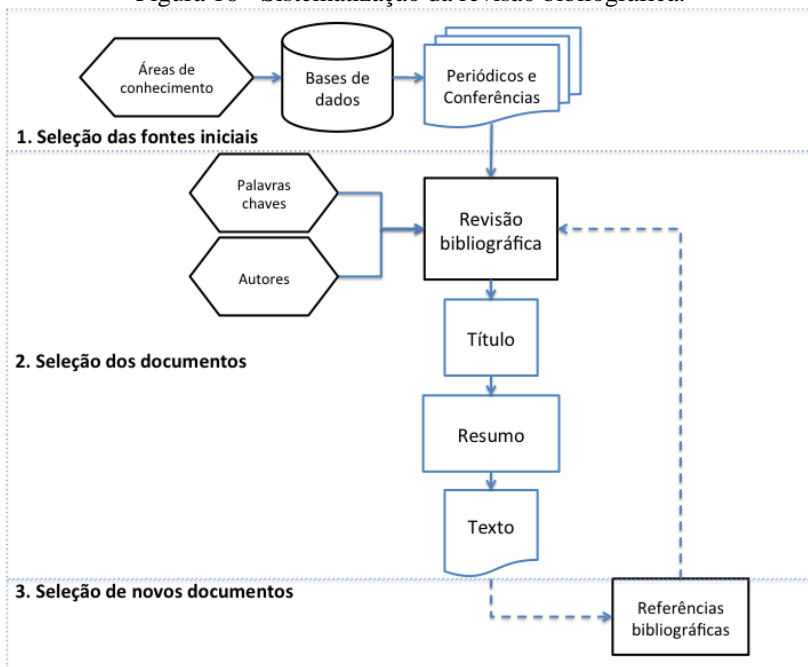
A sistematização utilizada está apresentada na Figura 16, sendo que a mesma consiste da seleção das bases de dados, seguindo com a seleção dos artigos alinhados ao foco da pesquisa.

Na seleção das bases de dados, fez-se primeiramente a escolha das áreas de conhecimento, as quais foram: multidisciplinar, engenharias e ciências sociais aplicadas. Com base nessas áreas de conhecimento e das bases disponíveis no portal de periódicos da CAPES (CAPES, 2013), definiu-se as seguintes bases de dados identificadas para a pesquisa:

- **Ciências Sociais Aplicadas** - INFORMS; SAGE Journals Online; Social Sciences Full-Text (Wilson); Humanities Abs & Full Text (Wilson).
- **Engenharias** - Emerald Fulltext (Emerald); IEEE Xplore; Compendex (Engineering Village 2).
- **Multidisciplinares** - Oxford Journals (Oxford University Press); SCOPUS; ScienceDirect (Elsevier); Cambridge Scientific Abstracts - CSA; Gale - Academic OneFile.

O resultado da pesquisa foi a não identificação da existência de trabalhos que utilizavam a TRI como método de *benchmarking*, conforme proposta pela pesquisa aqui apresentada.

Figura 16 - Sistematização da revisão bibliográfica.



Fonte: Villas, Macedo-Soares e Russo (2008), adaptado.

Na seleção dos documentos, utilizou-se de dois critérios básicos, primeiramente a definição do filtro com palavras chaves e como complemento, a conferência se houve a seleção de artigos de autores referências nas áreas da pesquisa. O filtro de palavras chaves utilizados foi feito em duas etapas (primeiramente em inglês e também em português):

- 1º filtro - [benchmarking **AND** "quantitative technique"].
- 2º filtro - [benchmarking **AND** "quantitative technique"] **AND** "case study".

A busca de artigos utilizou-se da ferramenta de metabusca do portal de periódicos da CAPES com o objetivo de identificar, com a análise dos títulos e resumos, quais documentos eram relevantes para o desenvolvimento da tese, conforme mostrado na Figura 17. A realimentação (linha pontilhada da Figura 16) da terceira etapa do processo na seleção de novos documentos serviu para refinar a busca com base nas referências dos documentos selecionados primeiramente.

Com base na Figura 3, o recorte temporal da pesquisa será a partir da 2ª geração do *benchmarking*, ou seja, a partir do final da década de 1970.

Figura 17 - Estratégias de buscas no Portal periódicos CAPES.

Portal periódicos CAPES

MAPA DO SITE | FALE CONOSCO | MEU ESPAÇO | TAMANHO A+ A- | CONTRASTE A+ A-

PÁGINA INICIAL | INSTITUCIONAL | ACERVO | BUSCA | NOTÍCIAS | SUPORTE

Você está aqui: Página inicial

Meu Espaço

Meus Artigos | Meus Conjuntos de Bases | Meus Periódicos | Minhas Estratégias de Busca | Meus Certificados | Ajuda

Minhas Estratégias de Busca

Clique no Nome da busca para executar a mesma:

No.	Busca	Bases	Alerta	Intervalo	Executado em	Ações
1.	Todos os campos-(benchmarking) E Todos os campos-(quantitative technique)	11 Bases.				🗑️ 🔔 🖨️
2.	Todos os campos-(benchmarking) E Todos os campos-(quantitative method)	11 Bases.				🗑️ 🔔 🖨️
3.	Todos os campos-(benchmarking) E Todos os campos-(quantitative methodology)	11 Bases.				🗑️ 🔔 🖨️
4.	Todos os campos-(benchmarking) E Todos os campos-(quantitative technique) E Todos os campos-(case study)	11 Bases.				🗑️ 🔔 🖨️
5.	Todos os campos-(benchmarking) E Todos os campos-(quantitative methodology) E Todos os campos-(case study)	11 Bases.				🗑️ 🔔 🖨️
6.	Todos os campos-(benchmarking) E Todos os campos-(quantitative method) E Todos os campos-(case study)	11 Bases.				🗑️ 🔔 🖨️

Fonte: Elaborado pelo Autor (2011).

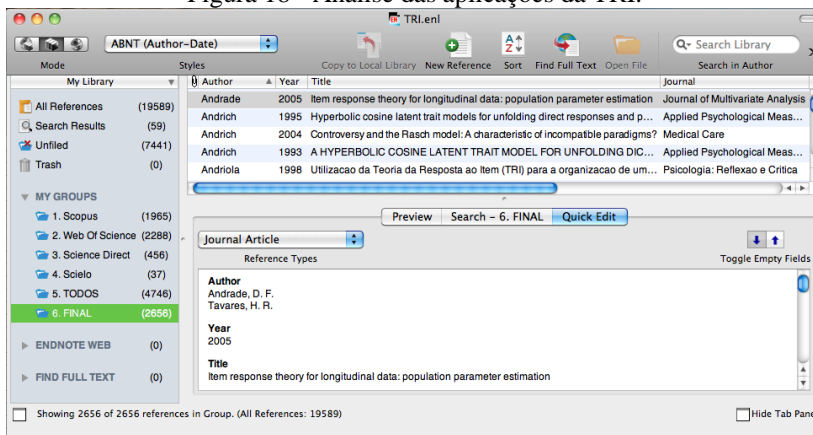
Na aplicação do 1º filtro nas bases selecionadas foram selecionados 176 artigos. Na aplicação do 2º filtro nas bases selecionadas foram selecionados 108 artigos, que foram oriundos do 1º filtro. Dos artigos selecionados pelo 1º filtro somente 11 (6,25%) eram aplicações de *benchmarking* utilizando técnicas quantitativas. Dos artigos do 2º filtro somente 6 (3,41%) eram aplicações de *benchmarking* de estudos de casos utilizando técnicas quantitativas. Em nenhum dos artigos resultantes dos filtros utilizou-se a TRI como técnica quantitativa de *benchmarking*.

Para corroborar com o ineditismo da tese, fez-se uma análise específica de trabalhos realizados com a aplicação da TRI. A análise desenvolvida baseou-se na identificação das áreas de conhecimento em que a TRI é usualmente aplicada. A pesquisa foi feita nas bases: SCOPUS, WEB OF SCIENCE, SCIENCE DIRECT e SCIELO.

O filtro utilizado foi [“item response theory” **OR** “Teoria da resposta ao item”], sendo que foram pesquisados somente artigos em periódicos (*journals*) e não foi feita limitação quanto à data da publicação.

Para a organização e quantificação dos resultados da pesquisa utilizou-se da ferramenta EndNote (THOMSON REUTERS, 2011). Inicialmente a pesquisa nas quatro bases gerou 4.746 artigos, que após um processo de organização para retirar duplicados chegou-se em 2.256 artigos, conforme pode ser visto na tela final das pesquisas na Figura 18.

Figura 18 - Análise das aplicações da TRI.



Fonte: Elaborado pelo Autor (2011).

Para identificar as áreas de conhecimento, fez-se a classificação dos periódicos nos quais os artigos foram encontrados pelas áreas de conhecimento, resultando nos números da Tabela 1.

Tabela 1 - A TRI aplicada nas diversas áreas de conhecimento.

Área do Conhecimento	Total	% Total
Psicologia	1.093	41%
Medicina	586	22%
Educação	473	18%
Multidisciplinar	305	11%
Estatística	83	3%
Ciências Sociais	47	2%
Administração	35	1%
Ciência da Computação	29	1%
Engenharia	5	0%
Total	2.656	100%

Fonte: Elaborado pelo Autor (2011).

Com base na Tabela 1, percebe-se a concentração dos artigos da TRI em 3 áreas do conhecimento: Psicologia, Educação e Medicina, perfazendo um total de 81%. Também é percebido uma quantidade incipiente de aplicações nas áreas de: Ciências Sociais, Administração, Ciência da Computação e Engenharia, com menos de 5% do total de artigos. É possível ainda refinar esses valores com uma análise mais criteriosa da categoria Multidisciplinar. O resultado desta segunda pesquisa foi também a não identificação da existência de trabalhos que utilizavam a TRI como método de *benchmarking*.

Com base nestes resultados, também verifica-se a pouca utilização da técnica nas áreas de conhecimento da Engenharia, reforçando a lacuna da aplicação da técnica nessa área.

5 - MÉTODO BIRTH

O objetivo deste capítulo é apresentar o processo teórico do novo método de *benchmarking*, denominado de BIRTH (*Benchmarking by Item Response Theory*). O mesmo está estruturado da seguinte forma: primeiramente são apresentados os princípios do novo método, em seguida são desenvolvidas as etapas de elaboração do teste e da definição dos itens, apresentada a proposta de construção da escala de competitividade e finalmente são estimados os índices dessa escala.

5.1 A ADAPTAÇÃO DO PROCESSO

O princípio do novo método é avaliar a competitividade de um produto, modelando-a como um traço latente. Essa é a essência do método de *benchmarking* proposto neste trabalho, em que na impossibilidade de uma medida direta da competitividade, esta será medida por intermédio da TRI.

A base para o desenvolvimento do novo método de *benchmarking* consiste do processo de construção de escalas psicométricas (Figura 6). Com a apropriação dos princípios de medição de um traço latente, a proposta está pautada na utilização de etapas do processo de construção das escalas com a TRI, substituindo as etapas do processo clássico de *benchmarking* apresentado na Figura 5. Entende-se que o método proposto possibilita o desenvolvimento de todos tipos de aplicações de *benchmarking* apresentados em 2.3.

A inserção é proposta com base na equivalência dos objetivos das etapas que perfazem todo processo de construção de uma escala psicométrica com o intuito das etapas iniciais do processo clássico de *benchmarking*. Na etapa denominada Teoria, ocorre a definição do traço latente competitividade, com as propriedades e sua dimensionalidade. Na etapa Construção, o intuito é definir, operacionalizar e obter um estudo piloto dos itens, que são baseados nas características dos produtos. Em Validação, realiza-se o planejamento, a aplicação do instrumento de medição sobre os produtos, e a análise da dimensionalidade, dos itens e da precisão da escala, com o uso dos modelos estatísticos da TRI. E, finalmente, na Normalização, são definidas as normas da escala competitividade. Essas etapas descritas serão denominadas simplesmente de TRI e o uso destas tem o intuito de substituir as etapas de Planejamento e Análise do processo de *benchmarking*, proporcionando, assim, a criação do novo método. As

demais etapas desse processo (Integração, Ação e Maturidade) serão executadas conforme modelo clássico da Figura 5.

A proposta de equivalência das etapas de *benchmarking*, que correspondem aos cinco primeiros passos do Planejamento e Análise, com as etapas de construção de uma escala psicométrica, estão apresentadas no Quadro 5.

Quadro 5 - Equivalência proposta das etapas *benchmarking* com as etapas de construção de uma escala psicométrica.

Benchmarking		Escala Psicométrica
Fases	Passos	Passos
Planejamento	P1. Identificar o que vai ser comparado	P1. Sistema psicológico P2. Propriedade P3. Dimensionalidade
	P2. Identificar as empresas referências	P4. Definições
	P3. Determinar os dados e o método de coleta desses	P5. Operacionalização P6. Análise piloto dos itens P7. Planejamento da aplicação P8. Aplicação e coleta
Análise	P4. Determinar a diferença do desempenho atual	P9. Dimensionalidade P10. Análise dos itens P11. Precisão das escalas
	P5. Projetar o nível de desempenho futuro	P12. Estabelecimento de normas

Fonte: Elaborado pelo Autor (2011).

A proposta de processo do novo método de *benchmarking* (BIRTH) está representado no fluxograma da Figura 19.

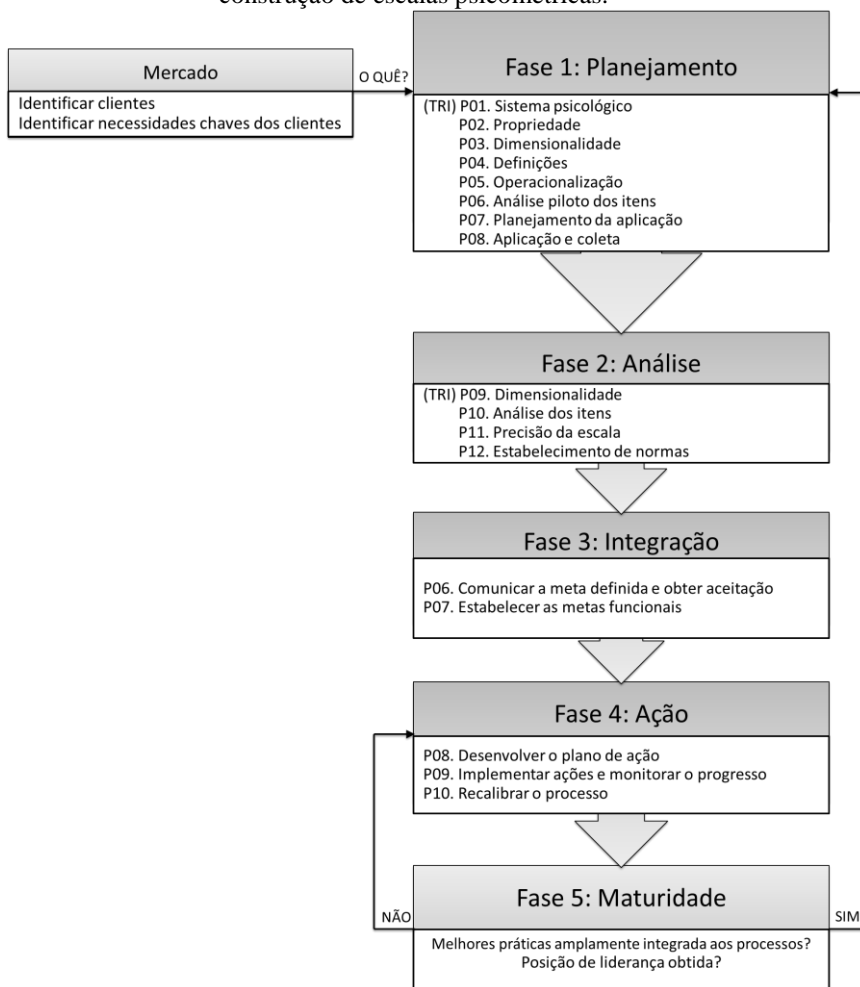
5.2 O CONSTRUTO, O TESTE E OS ITENS

O novo método inicia com a definição do traço latente da competitividade dos produtos (bens ou serviços) das empresas que serão comparados, e o levantamento dos dados destes tem que seguir a metodologia de construção de itens de uma escala psicométrica.

A definição dos itens é uma das etapas mais importantes para o desenvolvimento do processo BIRTH, pois tem, essencialmente, o intuito de modelar a competitividade dos produtos, com base nas características desses no mercado. As fontes dos itens podem ser

oriundas da literatura ou de entrevistas com especialistas (Figura 6), ou mesmo dos próprios clientes (Figura 5).

Figura 19 - Proposta do novo método de *benchmarking* com base na construção de escalas psicométricas.



Fonte: Elaborado pelo Autor (2011).

Para exemplificar a construção dos itens, no Quadro 6 estão descritos modelos de como construir três tipos básicos de itens, conforme os tipos de variáveis (binária, ordinal e intervalar), que podem

representar as características de um produto. Os itens exemplos representam uma característica genérica, denominada simplesmente de *X*.

Quadro 6 - Itens propostos por tipo de variável.

Tipo de Variável	Tipo de Item	Exemplo de Item	Exemplos de Respostas		
Binária	Dicotômico	Possui característica <i>X</i> ? ^a	Sim		Não
Ordinal	Politômico	Qual o nível da característica <i>X</i> ?	Ordem 1	Ordem 2	Ordem 3
Intervalar	Politômico	Qual o valor da característica <i>X</i> ?	Valor		

Fonte: Elaborado pelo Autor (2011).

Com base no Quadro 6, no caso de um item do tipo dicotômico binário, o objetivo é verificar se o produto possui ou não uma determinada característica. No caso de um item politômico ordinal, busca-se saber a intensidade ou a ordem de complexidade de uma característica. E finalmente, em um item politômico intervalar tem-se o valor contínuo da característica do produto, sendo que esse último exemplo de variável, pode ser transformada em uma variável binária ou ordinal, por meio da definição de categorias.

No Quadro 7, Quadro 8 e Quadro 9, seguem alguns exemplos de características de produtos e respectivos itens que poderiam representá-los, bem como o tipo de variável que o item representa. Os exemplos são de produtos e serviços de informática, em que foram apresentadas duas características tangíveis (físicas) de um produto: tela e sistema operacional; e, também, duas características intangíveis (serviço): acesso à Internet e serviço de suporte.

Quadro 7 - Exemplos de itens de uma variável binária, de produtos e serviços de informática.

Exemplo de item	Exemplo de Respostas	
Possui tela plana?	Sim	Não
Possui mais de um sistema operacional?	Sim	Não
Tem acesso à Internet?	Sim	Não
Possui serviço de suporte?	Sim	Não

Fonte: Elaborado pelo Autor (2011).

Quadro 8 - Exemplos de itens de uma variável ordinal, de produtos e serviços de informática.

Exemplo de item	Exemplo de Respostas		
Qual a importância de ter uma tela plana?	Baixa	Média	Alta
Qual a importância de ter mais de um sistema operacional?	Baixa	Média	Alta
Qual a satisfação com o acesso à Internet?	Insatisfeito	Satisfeito	Muito Satisfeito
Qual a satisfação com o serviço de suporte?	Insatisfeito	Satisfeito	Muito Satisfeito

Fonte: Elaborado pelo Autor (2011).

Quadro 9 - Exemplos de itens de uma variável intervalar, de produtos e serviços de informática.

Exemplo de item	Exemplo de Respostas
Qual a resolução da tela?	resolução em pixels: - 307.200 pixels (640x480) - 480.000 pixels (800x600) - 786.432 pixels (1024x768)
Possui quantos sistemas operacionais (SO)?	quantidade de SO: - 1 SO - 2 SO - 3 SO
Qual a velocidade de acesso à Internet?	velocidade em bits por segundo (bps): - 256 Kbps - 2 Mbps - 10 Mbps
Quantas horas de atendimento do serviço de suporte por dia?	horas de atendimento diárias (h): - 8 h - 16 h - 24 h

Fonte: Elaborado pelo Autor (2011).

Com os exemplos apresentados, de uma mesma característica é possível construir-se diferentes tipos de itens, como por exemplo a tela de um produto eletrônico. Essa pode ser modelada por um item binário, se a mesma é plana ou não; por um item ordinal, como a importância de ter uma tela colorida (baixa, média ou alta); por um item intervalar (quantitativo), como a resolução da tela em pixels (pontos). Essa mesma

análise vale para a característica sistema operacional (SO) de um produto de informática.

Assim como para as características tangíveis, as características intangíveis podem ser modeladas por itens, como por exemplo, o serviço de suporte ao atendimento de clientes. Esse pode ser modelado por um item binário, se o serviço possui ou não suporte; por um item ordinal, como qual a satisfação com o serviço de suporte (insatisfeito, satisfeito ou muito satisfeito); por um item intervalar (quantitativo), como o número de horas diárias do serviço de suporte. Essa mesma análise vale para a característica acesso à Internet.

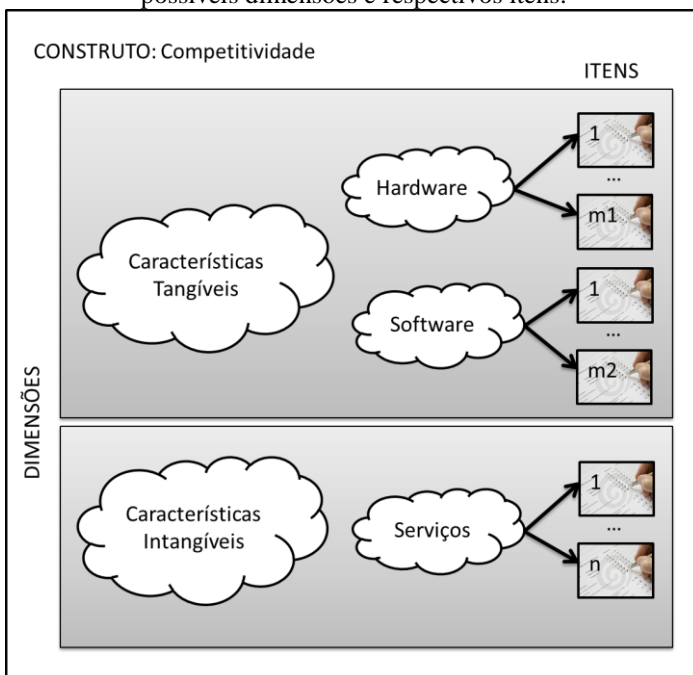
Assim, deve-se levar em consideração o tipo de item (binário, ordinal ou intervalar), atrelado a definição da dimensionalidade, que será utilizada para modelar a competitividade de um produto, pois dependendo destes, deve-se escolher o modelo da TRI apropriado na modelagem da competitividade. Alguns desses modelos foram apresentados nas expressões de (3.1) a (3.8).

Outro ponto que deve ser levado em consideração, na definição dos itens de um produto, é a dimensionalidade da competitividade, pois alguns produtos podem possuir mais de uma dimensão na modelagem da sua competitividade. Na sua essência, um produto tem uma dimensão tangível (física), geralmente ligada às características técnicas e tecnológicas destes, e em alguns casos, uma dimensão intangível, ligada a serviços que a empresa presta aos clientes do produto. Como os exemplos apresentados do Quadro 7 ao Quadro 9, das características tangíveis, no caso de um produto de informática, pode-se ter uma dimensão *hardware* e outra *software*. Das características intangíveis, pode-se ter uma dimensão ligada ao atendimento comercial de novos clientes e outro ao suporte dos clientes atuais.

Ao final desta etapa, com a definição da dimensionalidade e dos respectivos itens, tem-se o construto da competitividade do produto, representado, esquematicamente com base nos exemplos, na Figura 20.

Um ponto importante na definição dos itens, que não está relacionado com o número de dimensões, mas em construí-los de forma que exista independência entre os mesmos. E assim, ao final desse passo obtém-se o construto materializado em dimensões e itens.

Figura 20 - Exemplo de construto da competitividade de um produto, possíveis dimensões e respectivos itens.



Fonte: Elaborado pelo Autor (2011).

5.3 A APLICAÇÃO DOS MODELOS DA TRI

Com a elaboração do construto com suas dimensões e respectivos itens de um produto, o próximo passo é a definição do modelo estatístico da TRI mais adequado para representar a competitividade de um produto por meio da estimação dos parâmetros desse. Propõe-se que, os parâmetros dos modelos estatísticos da TRI, representem uma forma de modelar e, conseqüentemente, mensurar o traço latente da competitividade por intermédio de itens adequadamente construídos.

Assim, como na avaliação educacional, os modelos acumulativos são percebidos como os mais apropriados para modelar a competitividade, uma vez que os produtos, inseridos no mercado, “aprendem” e tendem a “responder” itens mais difíceis, que representam as melhores práticas de mercado, conforme aumentam o seu nível de

competitividade ao longo do tempo. Para exemplificar como um modelo estatístico da TRI pode representar o traço latente competitividade, foi utilizado o modelo logístico de dois parâmetros (ML2), apresentado em (3.2), pois parte-se do pressuposto que o parâmetro c (assíntota inferior) sempre será zero, pois é improvável um produto possuir uma característica ao acaso.

Estabelece-se uma relação que os produtos das empresas são equivalentes a alunos em uma avaliação educacional ou mesmo a pacientes em uma avaliação psicométrica, onde cada produto responde a um questionário composto por itens, um teste desenvolvido para mensurar o nível de competitividade deste. Logo a interpretação dos parâmetros de (3.2) para modelar a competitividade é a seguinte:

- U_{ij} indica a existência da característica (item) i no produto j , quando $U_{ij}=1$, o produto j tem a característica i , se $U_{ij}=0$, o produto j não tem a característica i ;
- θ_j é o índice de competitividade do produto j , representando o nível de competitividade desse produto no mercado;
- a_i é o parâmetro de inclinação do item i , representa o poder de discriminação dos produtos na característica i , é chamado de parâmetro de discriminação, onde espera-se que $a_i > 0$;
- b_i é o parâmetro de posição do item i , representa a dificuldade de se ter satisfatoriamente a característica i , é conhecido como parâmetro de dificuldade e está na mesma escala do índice de competitividade;
- $P(U_{ij} = 1/\theta_j)$ é a probabilidade do produto j , com nível de competitividade θ_j , possuir a característica i .

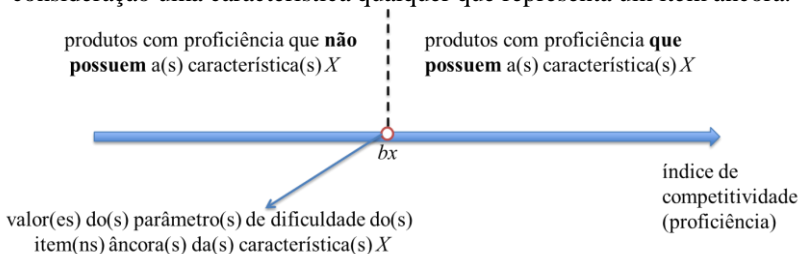
No modelo ML2, escala (0, 1), quanto maior o valor do parâmetro a , maior é a discriminação proporcionada pela característica representada pelo item. O valor de b representa a dificuldade do item, quanto maior o valor deste parâmetro, maior a dificuldade para ter a característica representada pelo item. Os valores de b geralmente estão no intervalo de ± 3 e o mesmo está na mesma unidade do índice de competitividade. Os parâmetros dos itens são estimados utilizando-se as respostas apresentadas e métodos de estimação, sendo que o mais utilizado é o método da máxima verossimilhança. Com a calibração dos parâmetros dos itens, faz-se a construção da escala de competitividade, tema da próxima seção.

5.4 A ESCALA DE COMPETITIVIDADE

Com os parâmetros dos itens estimados, chega-se no passo mais importante do método, que é a construção de uma escala de competitividade. Partindo-se da premissa teórica, que os parâmetros de dificuldade dos itens estão na mesma unidade de medida do índice de competitividade de um produto, o primeiro ponto a desenvolver é a identificação dos itens âncoras. Para a identificação destes, as condições (3.12), (3.13) e (3.14) devem ser satisfeitas. Na prática, caso as três não ocorram simultaneamente, pode-se flexibilizar e considerar somente duas, desde de que a terceira condição esteja próxima do valor pressuposto nas condições.

É por intermédio dos itens âncoras que é possível interpretar as proficiências estimadas dos produtos, criando grupos e identificando as melhores práticas de mercado. Por exemplo, dada a existência de um conjunto de itens âncoras relativos a características de produtos, que representam boas práticas deste. Considerando estimados os índices de competitividade desse produto, os que possuem o valor abaixo dos parâmetros de dificuldade dos itens âncoras, são produtos que não teriam as boas práticas implementadas. Entretanto, se o valor do índice for maior que os parâmetros de dificuldade dos itens âncoras, então o produto teria as boas práticas como características. De forma genérica, a lógica está exemplificada na Figura 21.

Figura 21 - Interpretação da escala de proficiência levando-se em consideração uma característica qualquer que representa um item âncora.



Fonte: Elaborado pelo Autor (2011).

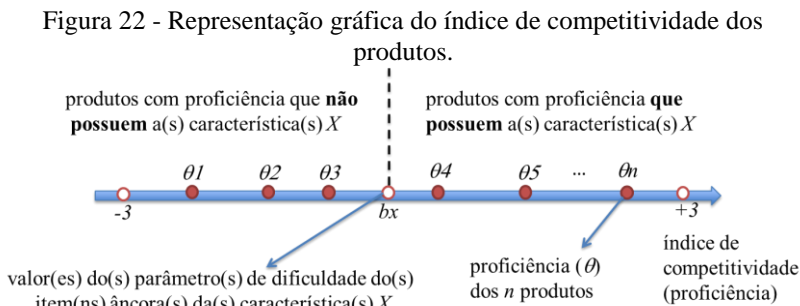
Apenas para exemplificar, dada a hipotética existência de um conjunto de itens âncoras, que representam a possibilidade de múltiplos sistemas operacionais e atualização destes por parte do usuário, ambos em um mesmo nível de dificuldade (bx). Os produtos com as melhores

práticas, cujos índices de competitividade estão acima destes itens âncoras (bi), interpreta-se que com relação a estas características é que esses possuam mais de um sistema operacional (multiplataformas) e com a possibilidade de atualização do *software* por parte do usuário. Considerando estimadas todas as proficiências de produtos, essa análise é repetida para cada item(ns) âncora(s) de forma a verificar os que possuem a proficiência acima dos valores dos parâmetros de dificuldade dos âncoras, ou seja, as melhores práticas e, portanto, um maior nível de competitividade no mercado.

5.5 O ÍNDICE DE COMPETITIVIDADE

Com a escala de competitividade construída, chega-se ao último passo, que é o cálculo das proficiências de cada produto, que corresponderá ao índice de competitividade destes. Para o cálculo dos mesmos utiliza-se um dos modelos estatísticos apresentados de (3.1) a (3.8), que foi utilizado na avaliação do produto, por meio dos passos apresentados na seção 3.4.2. Neste caso, propõe-se que, a proficiência seja a medida da competitividade dos produtos.

A medida da competitividade é estimada primeiramente na mesma escala na qual os itens foram calibrados, geralmente num intervalo de ± 3 . Quanto maior o valor do índice, maior tende a ser a competitividade do produto no seu mercado, pois significa que o produto apresentou a existência dos itens do construto. O resultado da estimativa dos índices pode ser representado graficamente em um eixo que corresponde à escala de competitividade, conforme apresentado na Figura 22.



Fonte: Elaborado pelo Autor (2011).

Com base na Figura 22, os produtos com proficiência θ_1 , θ_2 e θ_3 estão abaixo do item âncora bx , logo interpreta-se que estes não possuem as características X . Já os produtos com proficiência θ_4 ou superior possuem proficiência maior que a dificuldade do item âncora bx , logo entende-se que estes possuem a característica X .

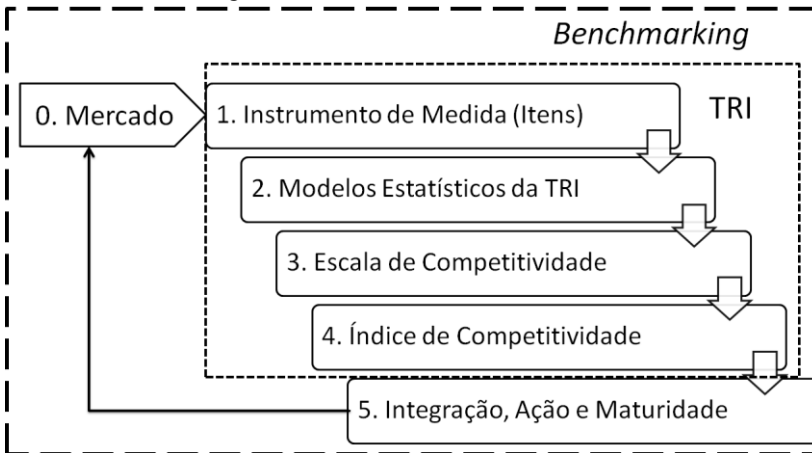
Apesar do índice ser calculado inicialmente na escala (0,1), o mesmo pode ser transformado para qualquer escala, com uma nova média e novo desvio padrão, por meio de uma transformação linear de escala. Assim como no seu uso em avaliações educacionais (BRASIL, 2009; OECD, 2012), em que o valor da proficiência identifica quais os conhecimentos que o aluno possui, pela verificação dos itens que possuem valores de dificuldades menores do que a proficiência do aluno, da mesma forma, o método proposto consegue identificar quais as melhores práticas que o produto possui. Além de identificar as boas práticas já existentes também é possível identificar as características prioritárias, de forma sequencial na escala construída, que devem ser implementadas, para aumentar o nível de competitividade do produto no mercado.

5.6 SÍNTESE DO MÉTODO

O novo processo de *benchmarking*, o método BIRTH (*Benchmarking by Item Response Theory*), tem como essência a proposta da substituição de etapas do seu processo clássico pelas etapas de aplicação da técnica da TRI, mostrando a relevância da avaliação de competitividade de produtos por meio de um traço latente. A Figura 23 sintetiza o método desenvolvido. Nesta figura, a TRI está representada nas etapas 1, 2, 3 e 4 como parte integrante do processo de *benchmarking*, que finaliza com a etapa 5 e realimenta a etapa 0, o mercado consumidor.

O método inicia com a competitividade modelada como um traço latente, podendo ser medida com o desenvolvimento de um teste, um instrumento de medida composto de itens, que representam essa competitividade no mercado. A elaboração do teste tem que levar em consideração, principalmente, a dimensionalidade e os tipos de itens que modelam a competitividade do produto. Já em suas primeiras etapas, o próprio processo de elaboração do teste, com base na literatura, em entrevistas com especialistas, ou mesmo com consumidores, promove o aumento e/ou atualização do conhecimento sobre o mercado do produto, sobre o qual se pretende inserir as melhores práticas.

Figura 23 - Síntese do método BIRTH.



Fonte: Elaborado pelo Autor (2014).

A etapa subsequente é a modelagem da competitividade pela TRI, com a definição de modelos estatísticos robustos e a avaliação dos parâmetros calibrados, que possibilitam que o teste, geralmente, composto por itens qualitativos, consiga mensurar, quantitativamente, o traço latente, com a estimativa das proficiências (o índice de competitividade) de cada produto. O processo da TRI elimina, por meio da análise da estimativa dos seus parâmetros, os itens não pertinentes ao traço latente modelado, a competitividade. Dessa forma, auxilia na consolidação do conhecimento dos itens desenvolvidos, e que realmente contribuem na modelagem da competitividade do produto em questão.

Na terceira etapa cria-se a escala de competitividade. Aqui encontra-se a principal contribuição da TRI no método proposto: oferecer a possibilidade de interpretação do índice de competitividade, por intermédio da identificação dos itens âncoras. O diferencial está na utilização do parâmetro de dificuldade dos itens, que representam o esforço de inserção de uma característica no produto, que está na mesma unidade de medida do índice de competitividade. Nesse contexto, a interpretação da escala sobre os itens do teste, possibilita aos gestores identificarem, primeiramente, a importância de uma determinada característica (item) e por conseguinte, a dificuldade de inserção dessa boa prática (item) no produto. Esta análise possibilita, então, a priorização das boas práticas alinhadas à competitividade do mercado.

Com a definição da escala de competitividade, a última etapa culmina na determinação dos índices de competitividade de cada

produto. A precisão do índice de competitividade, além da perfeita definição do teste, depende do número de itens utilizados e da precisão das estimativas dos parâmetros do modelo estatístico da TRI. Sendo ainda importante acrescentar, que a precisão da estimativa dos parâmetros depende do número de produtos submetidos ao teste. Esta análise consiste da identificação da posição (ranking) do produto no seu mercado.

Um ponto relevante do método no aspecto da forma de disponibilização dos resultados é a possibilidade de identificar sobre a escala de competitividade construída, atrelada à posição dos itens âncoras nesta escala, a visualização do índice de competitividade das empresas/produtos. A partir do valor do índice estimado, posiciona-se empresas/produtos na escala de competitividade calcada pela interpretação dos itens âncoras, definindo-se o conceito de empresas/produtos âncoras. Estes são interpretados e representam a materialização no mercado das empresas/produtos, explicados pelos seus respectivos itens âncoras.

Com a utilização do método BIRTH, o gestor pode, então, indicar qual boa prática deve ser inserida, prioritariamente, no seu produto, com o intuito de aumentar a sua competitividade, bem como conhecer a sua posição no mercado. E assim, volta-se ao processo clássico do *benchmarking*, nas fases de integração, ação e maturidade das novas práticas priorizadas e inseridas pelo gestor no seu produto.

6 - MÉTODO BIRTH APLICADO NA EDUCAÇÃO SUPERIOR

O intuito desse capítulo é apresentar uma aplicação prática do método BIRTH, com a construção de uma escala de competitividade para a comparação de um produto do tipo serviço. Escolheu-se o produto Educação, visto a disponibilidade e diversidade de dados primários de pesquisas desenvolvidas pelo MEC/INEP.

6.1 APRESENTAÇÃO DO PRODUTO EDUCAÇÃO SUPERIOR

O produto avaliado consiste da prestação de serviços educacionais, em que as informações das características desse produto responderam aos itens desenvolvidos no teste. Para as respostas ao construto elaborado, foram utilizadas fontes de dados primários e oficiais das características desses produtos. Dessa forma, tende-se a manter a homogeneidade e equidade das respostas, permitindo a comparabilidade dos resultados e garantindo a inexistência da influência de juízes no processo.

Para confirmar as premissas da fonte de informação utilizada, buscou-se a missão do INEP no Brasil. Este é o órgão responsável pelas estatísticas educacionais das diversas modalidades de ensino, sendo uma autarquia federal do MEC, e tem a missão de:

...promover estudos, pesquisas e avaliações sobre o Sistema Educacional Brasileiro com o objetivo de subsidiar a formulação e implementação de políticas públicas para a área educacional a partir de parâmetros de qualidade e equidade, bem como produzir informações claras e confiáveis aos gestores, pesquisadores, educadores e público em geral. Para gerar seus dados e estudos educacionais o Inep realiza levantamentos estatísticos e avaliativos em todos os níveis e modalidades de ensino, promove encontros para discutir os temas educacionais e disponibiliza também outras fontes de consulta sobre educação (BRASIL, 2012d).

As principais iniciativas desenvolvidas pelo INEP, que geram os dados oficiais sobre a educação brasileira, nas modalidades de educação infantil, ensino fundamental, ensino médio, educação especial, educação

de jovens e adultos (EJA) e da educação superior, estão apresentadas no Quadro 10.

Quadro 10 - Principais iniciativas desenvolvidas pelo MEC/INEP.

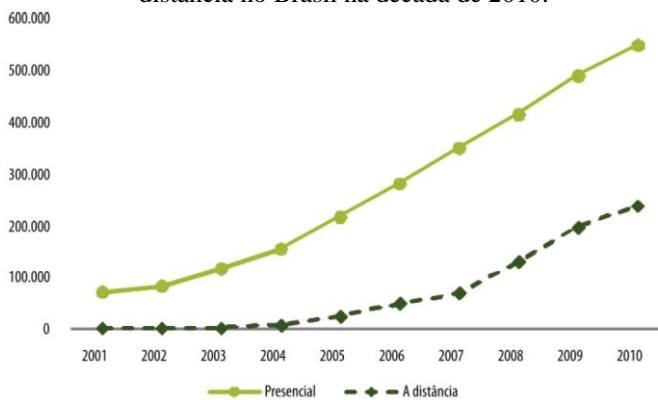
Iniciativa	Descrição
11. Censo Escolar da Educação Básica	levantamento anual de dados das modalidades de ensino regular (educação infantil e ensinos fundamental e médio), educação especial e educação de jovens e adultos das escolas públicas e privadas.
12. Sistema Nacional de Avaliação da Educação Básica (Saeb)	composto por duas avaliações bianuais complementares em Língua Portuguesa e Matemática: a Aneb (Avaliação Nacional da Educação Básica), que abrange de maneira amostral os estudantes do 5º e 9º anos do ensino fundamental e do 3º ano do ensino médio; a Anresc (Avaliação Nacional do Rendimento Escolar), aplicada em todos os alunos de 5º e 9º anos do ensino fundamental público, recebendo o nome de Prova Brasil
13. Provinha Brasil	avaliação da alfabetização (leitura e escrita) dos alunos do segundo ano de escolarização das escolas públicas brasileiras
14. Exame Nacional do Ensino Médio (Enem)	avaliação anual, criada em 1998, com o objetivo de avaliar o desempenho do estudante ao fim da educação básica, sendo que a partir de 2009 passou a ser utilizado também como forma de seleção para o ingresso no ensino superior
15. Exame Nacional de Certificação de Competências de Jovens e Adultos (Encceja)	avaliação voluntária e gratuita ofertada às pessoas que não concluíram os estudos em idade apropriada, que pode ser utilizada para a certificação do ensino fundamental, sendo que a certificação do ensino médio, para os maiores de 18 anos, é realizada pelo Enem.
16. <i>Programme for International Student Assessment</i> (PISA)	Programa Internacional de Avaliação de Estudantes, iniciativa internacional de avaliação comparativa da escolaridade básica (Leitura, Matemática, Ciências) entre estudantes na faixa dos 15 anos
17. Censo da Educação Superior	coleta anual de dados sobre a educação superior das instituições de ensino superior com cursos presenciais ou a distância, nas diferentes formas de organização acadêmica
18. Sistema Nacional de Avaliação da Educação Superior (SINAES)	sistema de avaliação das instituições, dos cursos e do desempenho dos estudantes da educação superior brasileira, que possui diversos instrumentos avaliativos coordenados e supervisionados pela Comissão Nacional de Avaliação da Educação Superior (Conaes)
19. Exame Nacional de Desempenho de Estudantes (Enade)	avaliação, com ciclo máximo trienal, para aferição do rendimento dos alunos dos cursos de graduação em relação aos conteúdos programáticos, suas competências e habilidades, integrando o Sinaes, juntamente com a avaliação institucional e a avaliação dos cursos superiores

Fonte: BRASIL (2012d).

No Quadro 10, percebe-se iniciativas relacionadas ao levantamento de dados estatísticos das características dos serviços educacionais (I1 e I7). Existem as iniciativas que definem os mecanismos de avaliação da qualidade do ensino (I2 e I8). E as demais iniciativas são as próprias avaliações das diversas modalidades de ensino existentes (I3, I4, I5, I6 e I9).

Com o foco no ensino superior, segundo o Censo da Educação Superior de 2010, no Brasil na década de 2010, houve um crescimento considerável nessa modalidade de ensino. Em 2001 eram 1.378 instituições de ensino superior (IES), chegando em 2010, a 2.378 instituições, um aumento de mais de 70%. Em matrículas, o aumento foi de mais de 110%, chegando ao número de 6.379.299 matrículas (BRASIL, 2012a). Nesse período, a maior expansão proporcional ocorreu nas matrículas na educação superior, nas instituições que oferecem cursos superiores de tecnologia (CST), apontando, de certa forma, uma direção dos investimentos na educação profissional e tecnológica (EPT), em que as matrículas em 2001 eram de 69.797, e em 2010 chegam a 781.609, perfazendo um aumento de mais de 1.000% (mais de onze vezes) (BRASIL, 2012a). A evolução dos CSTs por modalidade presencial e à distância está apresentada na Figura 24.

Figura 24 - Crescimento do número de matrículas em CSTs presenciais e à distância no Brasil na década de 2010.



Fonte: BRASIL (2012a).

Esses números pressupõem um aumento na competição desses cursos, e desta forma, reforçam a importância de um método que possibilite conhecer as melhores práticas desse mercado, proposta principal desse trabalho. Com base neste contexto, escolheu-se os CSTs

como foco do estudo de caso da aplicação do método apresentado no capítulo anterior. Considerando o foco da aplicação do método BIRTH nos CSTs, foram utilizadas as iniciativas I7 e I8, descritas no Quadro 10, como principais fontes de informação sobre o ensino superior brasileiro, estas são o Censo da Educação Superior e o SINAES.

Visto que o Censo de Educação Superior tem como principal objetivo disponibilizar “informações para a sociedade em geral, incluindo gestores, pesquisadores, administração pública, além de organismos internacionais” (BRASIL, 2012a, p. 13), este, por princípio, é a principal fonte das características do produto escolhido.

Os dados primários da pesquisa podem ser acessados no formato de microdados (BRASIL, 2012c), sendo os arquivos e respectivas quantidades de variáveis, que compõe essa base de dados, apresentados no Quadro 11.

Quadro 11 - Arquivos constituintes dos microdados do Censo de Educação Superior.

Arquivos	No. Variáveis	Descrição
IES.txt	40	Dados estatísticos das IES
CURSO.txt	63	Dados estatísticos dos Cursos
DOCENTE.txt	48	Dados estatísticos de Docentes
ALUNO.txt	85	Dados estatísticos de Aluno
LOCAL_OFERTA_POLO.txt	07	Dados estatísticos de Local de Oferta

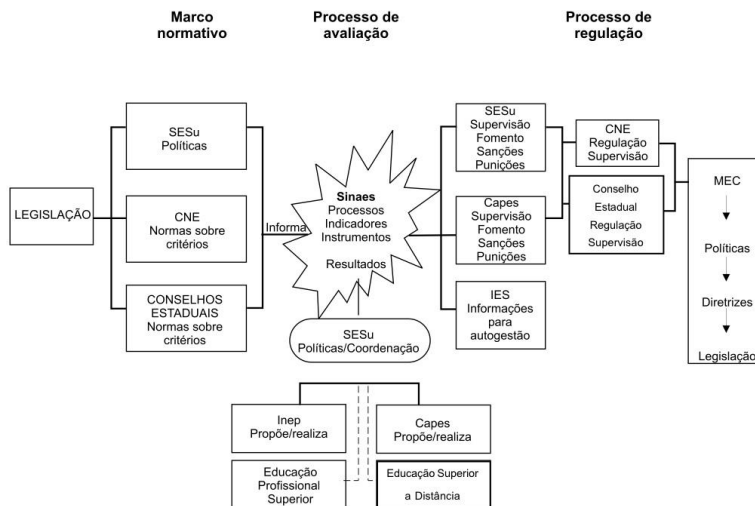
Fonte: BRASIL (2012c).

Os microdados, organizados em cinco arquivos, preservam o sigilo da identificação dos docentes e alunos, e o manuseio das informações somente é possível por meio de *softwares* de tratamento estatístico. Na importação dos dados, utilizou-se do pacote SPSS (*Statistical Package for the Social Sciences*) (SPSS, 2006) e dos *scripts* disponibilizados na documentação dos microdados (BRASIL, 2012c). A lista completa das variáveis, bem como os *scripts* utilizados na aplicação, estão transcritos nos Anexos deste trabalho.

Com relação à avaliação do ensino superior no Brasil, em 2004, foi instituído o SINAES, um sistema integrado de iniciativas com o objetivo de assegurar um processo nacional de avaliação de instituições, cursos e alunos (BRASIL, 2004). O SINAES é responsável por “articular, de forma coerente, concepções, objetivos, metodologias, práticas, agentes da comunidade acadêmica e de instâncias do governo” (BRASIL, 2009, p. 92). Na Figura 25 está representada a

responsabilidade do SINAES no processo de avaliação, com a responsabilidade sobre os processos, indicadores e instrumentos, bem como pelos resultados. O SINAES faz a conexão entre o marco normativo e os processos regulatórios, subsidiando os órgãos regulatórios com os resultados das avaliações para que a legislação seja cumprida.

Figura 25 - SINAES no processo de avaliação do ensino superior no Brasil.



Fonte: BRASIL (2009).

Em virtude da responsabilidade pelos instrumentos de avaliação da educação superior no Brasil, entende-se que o SINAES é a principal fonte de informação para a elaboração do teste, foco da próxima seção.

6.2 O CONSTRUTO, O TESTE E OS ITENS DO ENSINO SUPERIOR

A elaboração do teste, com a definição dos itens relacionados ao traço latente competitividade dos serviços da educação superior, é a essência do processo BIRTH. São os itens que representarão as características dos cursos e, conseqüentemente, a competitividade destes no mercado.

Como referencial teórico para a construção dos itens, utilizou-se como base o atual instrumento de avaliação dos cursos superiores no Brasil (BRASIL, 2012b), que faz parte do Sistema Nacional de Avaliação da Educação Superior (SINAES), que, além dos cursos, se baseia também na avaliação das instituições e alunos (BRASIL, 2009). O instrumento de avaliação é composto de três dimensões, divididos em 63 indicadores, conforme estrutura apresentada no Quadro 12.

Quadro 12 - Estrutura do instrumento de avaliação dos cursos superiores presencial e a distância

Dimensão	No. Indicadores	Fontes de consulta
1. Organização Didático-Pedagógica	22	Plano de Desenvolvimento Institucional, Projeto Pedagógico do Curso, Diretrizes Curriculares Nacionais e Formulário Eletrônico no e-MEC
2. Corpo Docente e Tutorial	20	Projeto Pedagógico do Curso, Formulário Eletrônico no e-MEC e Documentação Comprobatória
3. Infraestrutura	21	Projeto Pedagógico do Curso, Diretrizes Curriculares Nacionais, Formulário Eletrônico no e-MEC e Documentação Comprobatória

Fonte: BRASIL (2009).

Alguns indicadores são somente para cursos presenciais, outros somente para cursos a distância. Existem indicadores específicos para cursos como, por exemplo, Medicina, Direito e Licenciatura. No instrumento de avaliação, cada indicador recebe a atribuição de um conceito numa escala ordinal de 1 a 5, sendo que os critérios de atribuição dos conceitos estão representados no Quadro 13. Com a aplicação do instrumento de avaliação, calcula-se uma nota média que fica entre 1 e 5 em cada dimensão, sendo a nota final do curso, uma ponderação entre as três dimensões (BRASIL, 2009).

A aplicação será desenvolvida sobre a dimensão Corpo Docente e Tutorial, para a qual foi elaborado um teste para medir a competitividade dos cursos superiores, composto por dezesseis itens, apresentados no Quadro 14. Para a construção dos itens, utilizou-se dos 20 indicadores da dimensão corpo “Docente e Tutorial” como norteadores no processo, sendo que na inexistência de uma referência direta, buscou-se uma referência indireta. Apenas para exemplificar, o item I02 relativo à formação dos docentes em pós-graduação *stricto*

sensu foi construído com base no indicador “2.7 Titulação do corpo docente do curso” (Anexo A.1). Da mesma forma o item I03 relativo ao percentual de doutores tomou como referência o indicador “2.8 Titulação do corpo docente do curso - percentual de doutores”.

Quadro 13 - Critérios de atribuição dos respectivos indicadores da dimensão

Conceito	Descrição
1	Quando os indicadores da dimensão avaliada configuram um conceito NÃO EXISTENTE
2	Quando os indicadores da dimensão avaliada configuram um conceito INSUFICIENTE
3	Quando os indicadores da dimensão avaliada configuram um conceito SUFICIENTE
4	Quando os indicadores da dimensão avaliada configuram um conceito MUITO BOM/MUITO BEM
5	Quando os indicadores da dimensão avaliada configuram um conceito EXCELENTE

Fonte: BRASIL (2009).

Cada item será respondido pelos cursos, utilizando-se como fonte, os resultados do censo da educação superior de 2010, pesquisa desenvolvida pelo INEP/MEC, disponibilizada no formato de microdados (BRASIL, 2012c). Considerando o foco nos CSTs, após o processamento estatístico dos arquivos dos microdados, o censo levantou informações de 4.653 cursos de 905 IES, totalizando 213.142 docentes e 1.067.462 alunos nesta modalidade.

Quadro 14 - Teste da dimensão Corpo Docente e Tutorial e respectivos itens

Arquivo	Id	Item
IES	I101	O percentual de despesa com docentes > 40% ?
	I102	O percentual de docentes com pós <i>stricto sensu</i> > 50% ?
DOCENTES	I103	O percentual de docentes com doutorado > 20% ?
	I104	O percentual de docentes com dedicação integral ou parcial > 60% ?
	I105	O percentual de docentes com dedicação exclusiva > 20% ?
	I106	O percentual de docentes que atuam em extensão > 10% ?
	I107	O percentual de docentes que atuam em gestão > 20% ?
	I108	O percentual de docentes que atuam em pós à distância > 2% ?
	I109	O percentual de docentes que atuam em pós presencial > 10% ?
	I110	O percentual de docentes que atuam em pesquisa > 10% ?

ALUNOS	I111	O percentual de alunos com estágio > 50% ?
	I112	O percentual de alunos com atividades de extensão > 5% ?
	I113	O percentual de alunos com atividades de monitoria > 5% ?
	I114	O percentual de alunos com atividades de pesquisa > 5% ?
	I115	O percentual de alunos trancados < 5% ?
	I116	O percentual de alunos evadidos < 10% ?

Fonte: Elaborado pelo Autor (2012).

As respostas dos cursos aos itens, que por serem dicotômicos, geraram a constituição de uma base de dados com variáveis binárias. A existência da característica (atribuição do “1”) ocorre quando a resposta é “sim” ao item, do contrário, quando a resposta é “não”, considera-se a não existência da característica (atribuição de “0”). A cada resposta de um curso gera uma linha como exemplos apresentados na Tabela 2.

Tabela 2 - Exemplo de respostas dos cursos aos itens do teste.

CÓDIGO		ITENS															
IES	CURSO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
01	001	0	1	1	0	1	0	1	0	0	1	1	0	1	1	0	0
01	002	1	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1
02	001	1	1	0	0	1	0	0	0	1	0	1	1	1	0	0	0
03	001	0	0	0	0	1	0	0	0	1	1	1	1	1	1	0	0
03	002	0	0	0	0	1	1	1	0	0	0	1	1	1	1	0	0

Fonte: Elaborado pelo Autor (2012).

Com o objetivo de uma análise de conjunto de cursos com características semelhantes, dentro de um mesmo mercado de atuação, a tabela de dados, exemplificada na Tabela 2, foi elaborada somente com cursos presenciais das áreas de Ciências, Matemática e Computação, Engenharia, Produção e Construção, totalizando 1.581 CSTs.

Para o desenvolvimento da análise estatística dos dados foi utilizada a *R Language* (R DEVELOPMENT CORE TEAM, 2012), um *software* livre, por meio do pacote *irtoys* (PARTCHEV, 2012), desenvolve a análise da TRI. Foram utilizados também, os mecanismos de estimação do BILOG-MG (ZIMOWSKI *et al.*, 2003) especificamente na estimativa dos parâmetros do modelo estatístico da TRI. Nos Apêndices estão apresentados os *scripts* utilizados nas análises.

6.3 A APLICAÇÃO DO MODELO DA TRI NO ENSINO SUPERIOR

Com a definição do traço latente competitividade, bem como do teste e respectivos itens, passa-se para a definição do modelo estatístico mais adequado para representar a competitividade dos CSTs. Os parâmetros dos modelos estatísticos da TRI representam a competitividade na dimensão Corpo Docente e Tutorial. Entende-se que o mercado gera a tendência dos cursos a “aprenderem” e “responderem” itens mais difíceis, conforme aumentam o seu nível de competitividade ao longo do tempo. Por esse motivo foi utilizado um modelo acumulativo. O modelo estatístico da TRI mais adequado para esta aplicação é o modelo logístico de dois parâmetros (ML2), conforme interpretação apresentada na seção 5.3, estabelecendo-se a relação que os CSTs são equivalentes a alunos em uma avaliação educacional.

Para chegar a escala de competitividade, seguiu-se o processo de análise da TRI, proposto em 3.4. Nesta análise, o objetivo é a calibração dos parâmetros do modelo estatístico apresentados em (3.2) e da proficiência, que consiste da quantização do traço latente, ou seja, a competitividade dos cursos. Em um teste com qualidade, os itens devem possuir graus de discriminação (*a*) acima de “+0,75” e diferentes parâmetros de dificuldade (*b*) em toda a escala de competitividade.

Na análise dos 16 itens, nove apresentaram as características adequadas, considerados itens válidos, sendo que o resumo da análise está apresentado na Tabela 3 e Tabela 4.

Tabela 3 - Resultado da análise da distribuição dos itens válidos.

Itens	Cursos	Ocorrências	%Ocorrências
ITEM02	1.581	1.149	72,7%
ITEM03	1.581	577	36,5%
ITEM04	1.581	680	43,0%
ITEM05	1.581	305	19,3%
ITEM06	1.581	543	34,3%
ITEM07	1.581	452	28,6%
ITEM09	1.581	158	10,0%
ITEM10	1.581	629	39,8%
ITEM14	1.581	65	4,1%

Fonte: Elaborado pelo Autor (2014).

Na Tabela 3, destaca-se que o ITEM02, que corresponde à existência de mais de 50% de docentes com pós-graduação *stricto sensu* no curso, foi a característica com o maior percentual de ocorrência (72,7%). Já o ITEM14, que corresponde à existência de mais de 5% dos alunos envolvidos em pesquisa no curso, foi a característica com o menor percentual de ocorrência (4,1%).

Tabela 4 - Resultado da análise da TRI dos itens válidos.

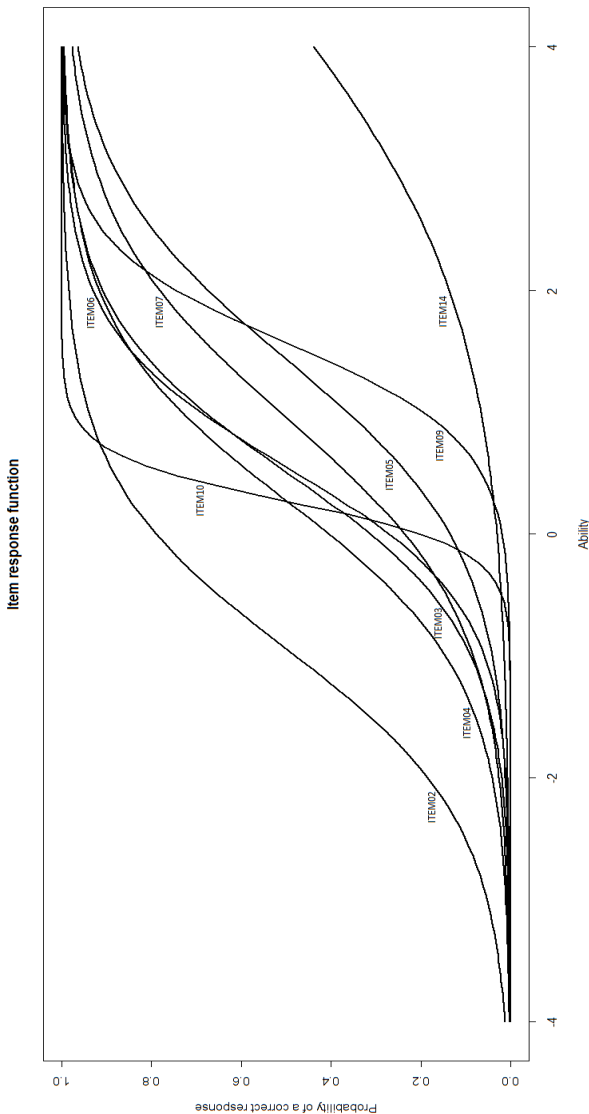
Itens	Discriminação (a)		Dificuldade (b)	
	estimativa	erro padrão	estimativa	erro padrão
ITEM02	1,405	0,085	-0,944	0,057
ITEM03	1,540	0,087	0,510	0,043
ITEM04	1,385	0,080	0,275	0,044
ITEM05	1,288	0,088	1,436	0,082
ITEM06	1,788	0,098	0,553	0,039
ITEM07	1,224	0,078	0,958	0,064
ITEM09	2,495	0,173	1,571	0,057
ITEM10	5,026	0,285	0,271	0,018
ITEM14	0,798	0,122	4,310	0,572

Fonte: Elaborado pelo Autor (2012).

Na Tabela 4, destaca-se que o ITEM10, que corresponde à existência de mais de 10% dos docentes envolvidos em pesquisa no curso, foi o maior parâmetro de discriminação ($a_{10}=5,026$). Apesar do valor ser muito alto, a decisão tomada como especialista em educação, entende-se que o envolvimento em pesquisa é um diferencial a ser considerado válido para o modelo, uma vez que os cursos superiores de tecnologia por terem um foco prático, ainda carecem de boas práticas como a pesquisa aplicada. Já o ITEM14, que corresponde à existência de mais de 5% dos alunos envolvidas em pesquisa no curso, foi o menor parâmetro de discriminação ($a_{14}=0,798$). Com relação ao parâmetro de dificuldade, o ITEM14 foi o item mais difícil ($b_{14}=4,310$ e o ITEM02, que corresponde à existência de mais de 50% de docentes com *stricto sensu* no curso, foi o item mais fácil ($b_2=-0,944$). A decisão de aceitar como válido o ITEM14, acima de +3, foi com o intuito de inserir um item referente aos ALUNOS, decisão também tomada como especialista em educação, para que o modelo ficasse mais próximo da realidade, pois, senão, dentro de uma mesma IES, todos os cursos teriam a mesma proficiência na dimensão Corpo Docente e Tutorial e esta condição

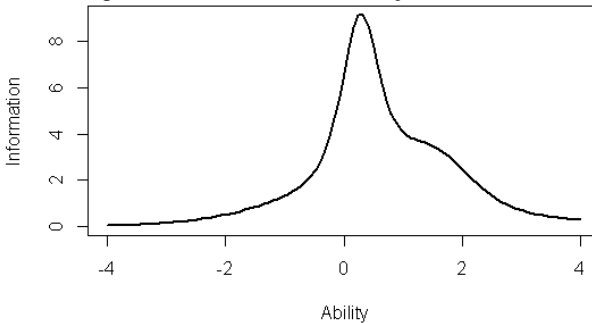
independência do desempenho/qualidade dos alunos. Na Figura 26 estão as curvas características dos nove itens válidos e a curva de informação do teste está apresentado na Figura 27.

Figura 26 - Curvas características dos itens válidos.



Fonte: Elaborado pelo Autor (2014).

Figura 27 - Curva de informação do teste.



Fonte: Elaborado pelo Autor (2012).

A curva de informação do teste apresenta que os itens possuem mais informação no intervalo de $[-0,5; 2,0]$. Idealmente, deve-se ter itens em toda a escala de competitividade.

6.4 A ESCALA DE COMPETITIVIDADE DO ENSINO SUPERIOR

Para a construção da escala, o processo começa com a identificação dos itens âncoras, conforme condições apresentadas em (3.12), (3.13) e (3.14). Na aplicação dessas três condições nos nove itens válidos, chegou-se a oito itens âncoras, conforme os valores de probabilidades destacados, calculados com o modelo ML2 estimado, apresentados na Tabela 5.

É na criação da escala que está a maior contribuição do método, pois esta possibilita interpretar a posição dos CSTs em função dos itens âncoras identificados. Os itens definidos no Quadro 14 foram condicionados às informações existentes nos microdados do Censo da Educação Superior referentes à dimensão Corpo Docente e Tutorial, sendo que somente nove dos dezesseis itens foram válidos para construção da escala de competitividade. Na Figura 28 é apresentada a escala de competitividade na forma gráfica, com os itens âncoras identificados.

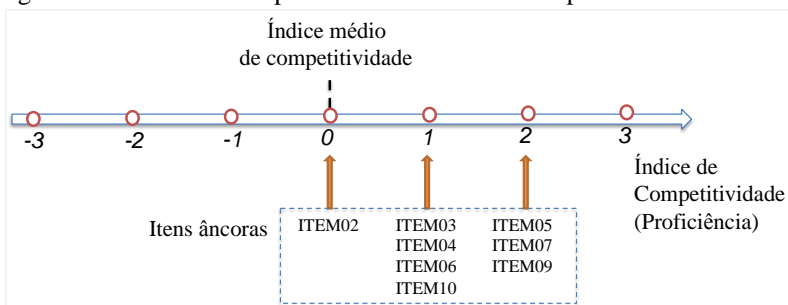
Tabela 5 - Determinação dos itens âncoras.

Item	Índice de Competitividade								
	-4,0	-3,0	-2,0	-1,0	0,0	1,0	2,0	3,0	4,0
ITEM02	0,01	0,05	0,18	0,48	0,79	0,94	0,98	1,00	1,00
ITEM03	0,00	0,00	0,02	0,09	0,31	0,68	0,91	0,98	1,00
ITEM04	0,00	0,01	0,04	0,15	0,41	0,73	0,92	0,98	0,99
ITEM05	0,00	0,00	0,01	0,04	0,14	0,36	0,67	0,88	0,96
ITEM06	0,00	0,00	0,01	0,06	0,27	0,69	0,93	0,99	1,00
ITEM07	0,00	0,01	0,03	0,08	0,24	0,51	0,78	0,92	0,98
ITEM09	0,00	0,00	0,00	0,00	0,02	0,19	0,74	0,97	1,00
ITEM10	0,00	0,00	0,00	0,00	0,20	0,98	1,00	1,00	1,00
ITEM14	0,00	0,00	0,01	0,01	0,03	0,07	0,14	0,26	0,44

Âncoras	ITEM02	ITEM03	ITEM05
		ITEM04	ITEM09
		ITEM06	ITEM07
		ITEM10	

Fonte: Elaborado pelo Autor (2012).

Figura 28 - Escala de competitividade dos CSTs e respectivos itens âncoras.



Fonte: Elaborado pelo Autor (2012).

A interpretação da escala de competitividade identificou a configuração de quatro grupos de CSTs com relação à dimensão Corpo Docente e Tutorial:

- não competitivos - índices abaixo de “0”, estão abaixo da média, e não possuem nenhuma das características dos itens âncoras;
- com baixa competitividade (pós-graduação *stricto sensu*) - índice no intervalo $[0, +1[$, definido pelo âncora ITEM02;

- com média competitividade (doutorado, dedicação integral/parcial, com atuação em pesquisa e extensão) - índice no intervalo [+1, +2[, definido pelos âncoras ITEM03, ITEM04, ITEM06 e ITEM10;
- com alta competitividade (dedicação exclusiva, atuação em gestão e em pós presencial) - índice de competitividade acima de +2, definido pelos âncoras ITEM05, ITEM07 e ITEM09.

Como foi utilizado um modelo acumulativo, entende-se que um grupo mais competitivo (maior índice de competitividade) possui, além das suas boas práticas com um diferencial, todas as boas práticas do grupo menos competitivo (menor índice de competitividade). Ressalta-se que os itens utilizados estão sintonizados com a ferramenta de avaliação utilizada pelo MEC e construídos conforme a disponibilidade das variáveis coletadas no Censo de Educação Superior.

6.5 O ÍNDICE DE COMPETITIVIDADE DO ENSINO SUPERIOR

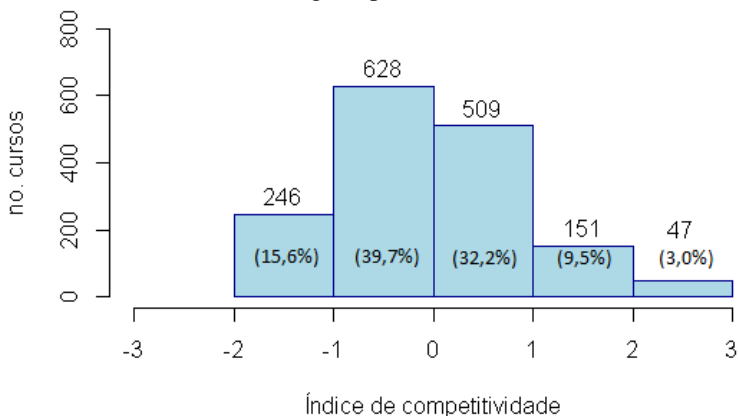
Com o modelo da TRI ajustado e a escala construída, o último passo foi estimar a proficiência de cada CST, que consiste do próprio índice de competitividade. Um valor que, com a TRI, é inicialmente mensurado com média nula e desvio padrão unitário.

Ao estimar a proficiência de um CST, é possível identificar a qual grupo, definido no item anterior, este pertence. Também existe a possibilidade do índice ser transformado para qualquer intervalo numérico por meio de uma mudança de escala, uma transformação linear. Como a dimensão Corpo Docente e Tutorial é medida pelo MEC por indicadores no intervalo de [1; 5], equivalente a uma média 3 e desvio padrão 1, no instrumento base da análise, é possível transformar o índice de competitividade para essa mesma escala.

A distribuição dos índices de competitividade dos CSTs estimados pelo método EAP, na escala original pelo modelo ML2, estão apresentados na Figura 29.

O valor “0” corresponde à média dos índices de competitividade dos cursos e quem está abaixo deste valor, encontra-se abaixo da média do cursos analisados. Com base na Figura 29, 874 (55,3%) dos CSTs estão no intervalo de [-2, 0], estando abaixo da média geral dos índices de competitividade. Já os demais CSTs (707; 44,7%) estão no intervalo de [0; +3], com o índice de competitividade acima da média.

Figura 29 - Distribuição dos índices de competitividade dos CSTs estimados na escala original pelo modelo ML2.



Fonte: Elaborado pelo Autor (2012).

A interpretação do índice de competitividade dos CSTs na escala construída, com base na distribuição dos dados da Figura 29, é feita da seguinte forma:

- CSTs não competitivos, abaixo da média, e não possuem nenhuma das características dos itens âncoras, sendo que nestas condições foram encontrados 874 cursos (55,3%);
- CSTs com baixa competitividade, que possuem mais de 50% dos docentes com pós *stricto sensu* (ITEM02), sendo que nestas condições estão 509 cursos (32,2%);
- CSTs com média competitividade, que se diferenciam dos cursos do grupo anterior por possuírem mais de 20% dos docentes com Doutorado (ITEM03), mais de 60% com Dedicção Integral / Parcial (ITEM04), mais de 10% com atuação em Extensão (ITEM06) e Pesquisa (ITEM10), perfazendo um total de 151 cursos (9,5%);
- CSTs com alta competitividade, que se diferenciam dos demais por possuírem mais de 20% dos docentes com Dedicção Exclusiva (ITEM05) e com atuação em Gestão (ITEM07) e mais de 10% atuando em Pós presencial (ITEM09), o que foi encontrado em 47 cursos (3,0%).

6.6 ANÁLISE DOS RESULTADOS

Um ponto importante da aplicação foi o uso de uma fonte homogênea de dados dos produtos, o Censo da Educação Superior. A importância desse aspecto é um dos suportes na aplicação do método, minimizando a influência de variáveis externas, que podem alterar a robustez do índice de competitividade.

Os resultados da aplicação do método na avaliação dos CSTs existentes em 2010, no Brasil, possibilitou validar a proposta desenvolvida, uma vez que foi criada uma escala de competitividade interpretativa e alinhada ao instrumento de avaliação vigente e aplica do ensino superior no Brasil. Percebe-se que o instrumento de avaliação na dimensão Corpo Docente e Tutorial (Anexo A.1) preconiza a formação dos docentes em pós-graduações *stricto sensu* e a experiência em docência com o foco em dedicação exclusiva ao ensino e as diversas formas de atividades ligadas à extensão, pesquisa e gestão.

Os resultados demonstraram que a análise dos itens âncoras, durante a construção da escala de competitividade, apontou quais as prioridades de foco de investimentos dos CSTs na dimensão Corpo Docente e Tutorial para que um curso aumente a sua competitividade no mercado. É possível, mas não foi foco desta aplicação, a identificação dos cursos que representam os níveis de competitividade definidos pelos itens âncoras, Estes cursos “âncoras” seriam os representantes no mercado dos quatro grupos de cursos identificados pela escala construída: os não competitivos e os de baixa, média, alta competitividade.

Entende-se que primeiramente os investimentos devem ser direcionados na formação dos docentes com pós-graduação *stricto sensu*, com o incentivo ao doutorado e que a dedicação seja preferencialmente exclusiva. A atuação dos docentes em pesquisa e extensão são pontos que aumentam a competitividade dos cursos. Os mais competitivos promovem atuação em gestão e em docência em programas de pós-graduação presenciais. Essa seria a ordem priorizada de investimentos que um gestor de um curso de educação superior, no Brasil, tem que levar em consideração para manter-se competitivo no mercado.

7 - CONCLUSÕES E TRABALHOS FUTUROS

7.1 CONCLUSÕES

O tema competitividade continua atual e a longevidade das organizações, que atendem as demandas de mercado, necessitam utilizar-se de técnicas que permitam a tomada de decisão dos gestores com a minimização dos riscos. Uma das técnicas que atende a essa necessidade é o *benchmarking*, que compreende da busca e internalização contínua das melhores práticas do mercado. As empresas âncoras buscam e são referência no mercado para a melhoria contínua da competitividade.

O processo de *benchmarking* é amplo e customizável, permitindo que as organizações utilizem-se dessa técnica da forma mais adequada a sua estrutura organizacional. Entende-se que as técnicas quantitativas são versáteis e poderosas no tocante às possibilidades de interpretação quali e quanti e é sobre este aspecto que o problema de pesquisa foi desenvolvido.

Conclui-se que o método desenvolvido, o BIRTH, demonstrou como o processo da TRI, uma técnica quantitativa, contribui para tornar o *benchmarking* mais robusto, possibilitando resolver dois pontos necessários: a priorização das boas práticas e a posição (*ranking*) da organização junto ao seu mercado.

Da mesma forma, conclui-se que os objetivos propostos foram alcançados, pois um novo método de benchmarking foi desenvolvido, possibilitando a construção de escalas de competitividade. Foi apresentada a aderência das etapas da TRI às etapas do processo clássico de benchmarking e foi desenvolvida uma aplicação do novo método desenvolvido.

Um ponto crítico do método está na definição dos itens do teste em quantidades suficientes que representem a competitividade deste com precisão. Outro ponto importante e relacionado à precisão do índice de competitividade está no número de produtos necessários para a estimação dos parâmetros do modelo estatístico utilizado para a modelagem da competitividade.

Conclui-se que o método desenvolvido está sintonizado à última geração, “Benchmarking em Rede”, onde a formação de redes para compartilhamento de boas práticas tende a aumentar a competitividade das organizações no mercado. Este pode ser utilizado em qualquer tipo de aplicação de *benchmarking*, seja definido pelas natureza da

referência, do conteúdo e do propósito. O método é acadêmico na sua essência, mas também se adapta aos outros dois tipos: consultor, e organização.

Um ponto de destaque são as vantagens que a TRI apresentou quando comparada com as duas principais técnicas quantitativas identificadas pelo referencial teórico, a DEA e AHP. A TRI simultaneamente apresenta os resultados proporcionados pela DEA e AHP. Com relação à DEA proporciona a classificação (*ranking*) dos produtos, mas com a vantagem de proporcionar interpretações em conjunto com os itens construídos, uma vez que os índices de competitividade dos produtos estão na mesma unidade de medida que o parâmetro de dificuldade dos itens. Com relação à AHP proporciona a priorização da inserção de novas características, mas com a vantagem de não depender da subjetividade de julgadores e possibilitar interpretações sobre a escala de competitividade construída sob a ótica dos itens âncoras.

7.2 TRABALHOS FUTUROS

Como trabalhos futuros, vê-se a possibilidade de:

- ampliar a aplicação desenvolvida na avaliação de cursos de ensino superior para outras dimensões (por exemplo, infraestrutura), elaborando itens com base teórica e/ou de especialistas para medir a competitividade mais ampla de instituições de ensino;

- desenvolver uma aplicação com um produto tangível (bem material), como por exemplo *notebooks*, *smartphones*, que dada a quantidade de variedades de marcas e tipos viabilizam a aplicação do método.

- desenvolver uma aplicação com um produto multidimensional, como por exemplo uma dimensão tangível (bem material) e outra intangível (serviços), como por exemplo fabricantes de produtos de telecomunicações, que além das características dos produtos (*hardware* e/ou *software*), disponibilizam serviços de suporte, garantia e também, viabilizam a aplicação do método.

- desenvolver uma aplicação com produtos cuja competitividade seja modelada com itens politômicos, que possuem mais informações do que os itens dicotômicos;

- adaptar o método BIRTH para o uso de itens condicionais, que modelaria a competitividade de produtos com características específicas.

- adaptar o método BIRTH para o uso de modelos de desdobramentos, que modelaria o posicionamento das empresas e de seus produtos no mercado.

REFERÊNCIAS

- AHMED, P. K.; RAFIQ, M. Integrated benchmarking: a holistic examination of select techniques for benchmarking analysis. **Benchmarking for Quality Management and Technology**, v. 5, n. 3, p. 225-242, 1998.
- ALEXANDRE, J. W. C.; ANDRADE, D. F. de; VASCONCELOS, A. P.; ARAÚJO, A. M. S. Uma proposta de análise de um construto para a medição dos fatores críticos da gestão pela qualidade através da teoria da resposta ao item. **Gestão & Produção**, v. 9, n. 2, p. 129-141, 2002a.
- ALEXANDRE, J. W. C.; ANDRADE, D. F. de; VASCONCELOS, A. P.; ARAÚJO, A. M. S.; BATISTA, M. J. **Teoria de Resposta ao Item: Aplicação do Modelo de Escala Gradual na Gestão pela Qualidade**. XXII ENEGEP - Encontro Nacional de Engenharia de Produção, 2002b.
- ANAND, G.; KODALI, R. Benchmarking the benchmarking models. **Benchmarking: An International Journal**, v. 15, Iss: 3, p. 257-291, 2008.
- ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. C. **Teoria da Resposta ao Item: Conceitos e Aplicações**. SINAPE, 2000.
- ANDRICH, D. A rating formulation for ordered response categories. **Psychometrika**, 43, p. 561-573, 1978.
- ANDRONIKIDIS, A. The application of quality function deployment in service quality management. **The TQM Journal**, v. 21, n. 4, p. 319-333, 2009.
- ASQ. **The Global Voice of Quality**. Disponível em: <http://asq.org/sixsigma/2008/04/ahp-matrix-template.html?shl=087837>. Acesso em 02 fevereiro 2013.
- BABCOCK, B. G. E. Estimating a Noncompensatory IRT Model Using a Modified Metropolis Algorithm. **PhD dissertation**, University of Minnesota, 2009.
- BAKER, F. B.; KIM, S-H. **Item Response Theory: Parameter Estimation Techniques**. New York: Marcel Dekker, 2004.
- BAYLEY, S. Measuring customer satisfaction. **Evaluation Journal of Australasia**, v. 1, n. 1, p. 8-16, 2001.

BEATON, A. E.; ALLEN, N. L. Interpreting scales through scale anchoring. **Journal of Educational Statistics**, v. 17, p. 191-204, 1992.

BHUTTA, K. S. HUQ, F. Benchmarking - best practices: an integrated approach. **Benchmarking: An International Journal**, v. 6, n. 3, p. 254-268, 1999.

BOCK, R. D. A Brief History of Item Theory Response. **Educational Measurement: Issues and Practice**, v. 16, n. 4, p. 21-33, 1997.

BOCK, R. D.; ZIMOWSKI, M. F. Multiple Group IRT. in VAN DER LINDEN, W. J.; HAMBLETON, R. K. **Handbook of Modern Item Response Theory**. New York: Springer-Verlag, 1997.

BORTOLOTTI, S. L. V. Aplicação de um modelo de desdobramento da teoria da resposta ao item - TRI. **Dissertação (Mestrado)**, PPGEP/UFSC, 2003.

BORTOLOTTI, S. L. V. Resistência à mudança organizacional: medida de avaliação por meio da teoria da resposta ao item. **Tese (doutorado)** - PPGEP /UFSC, 2010.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Censo da educação superior: 2010 - resumo técnico**. Brasília: INEP, 2012a. Disponível em:

<http://portal.inep.gov.br/web/censo-da-educacao-superior/resumos-tecnicos>. Acesso em: 20 julho 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Instrumento de Avaliação de Cursos de Graduação presencial e a distância**. Brasília: INEP, 2012b. Disponível em: <http://portal.inep.gov.br/superior-condicoesdeensino-manuais>.

Acesso em: 08 julho 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Microdados do Censo de Educação Superior**. Brasília: INEP, 2012c. Disponível em:

ftp://ftp.inep.gov.br/microdados/micro_censo_escolar_2010.zip. Acesso em: 08 julho 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Institucional INEP**. Brasília: INEP, 2012d.

Disponível em: <http://portal.inep.gov.br/web/aceso-a-informacao/institucional>. Acesso em: 20 julho 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **SINAES - Sistema Nacional de Avaliação da Educação Superior: Da Concepção à Regulamentação**. Brasília: INEP, 5^a ed., 2009. Disponível em:

<http://www.publicacoes.inep.gov.br/detalhes.asp?pub=4389>. Acesso em: 08 julho 2012.

BRASIL. Lei Nº 10.861, de 14 de ABRIL de 2004. Publicada no Diário Oficial da União nº 72 em 15 de abril de 2004, seção 1, p. 3-4.

BÜYÜKÖZKAN, G. A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry. **Expert Systems with Applications** 39, p. 2341-2354, 2012.

BÜYÜKÖZKAN, G. Strategic analysis of healthcare service quality using fuzzy AHP methodology. **Expert Systems with Applications** 38, p. 9407-9424, 2011.

CAMP, R. C. **Benchmarking: the search for industry best practices that lead to superior performance**. Milwaukee: ASQC Quality, 1989.

CAPES. Portal Periódicos da CAPES, 2013. Disponível em: <http://www.periodicos.capes.gov.br>. Acesso em: 06 janeiro 2013.

CARNEVALLI, J. A. MIGUEL, P. C. Review, analysis and classification of the literature on QFD - Types of research, difficulties and benefits. **Int. J. Production Economics** 114, p. 737-754, 2008.

CARPINETTI, L. C. R.; MELO, A. M. de. What to benchmark?: A systematic approach and cases. **Benchmarking: An International Journal**, v. 9, n. 3, p. 244-255, 2002.

COLEMAN, M. J.; MATTHYSSE, S.; LEVY, D. L.; COOK, S.; LO, J. B. Y.; RUBIN, D. B.; HOLZMAN, P. S. Spatial and object working memory impairments in schizophrenia patients: a bayesian item-response theory analysis. **Journal of Abnormal Psychology**, 111, n. 3, p. 425-435, 2002.

COOMBS, C. H. **A theory of data**. New York: Wiley, 1964.

CHARNES, A. COOPER, W.W. RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research** 2, p. 429-444, 1978.

COOPER, W.W. SEIFORD, L.M. TONE, K. **Data Envelopment Analysis: A Comprehensive Text with Models, Applications,**

References and DEA-Solver Software. Kluwer Academic Publishers: Boston, 2000.

COOPER, W.W. SEIFORD, L.M. ZHUS, J. **Handbook on Data Envelopment Analysis.** Kluwer Academic Publisher: Boston, 2004.

CORTON, M. L. Benchmarking Central American water utilities. **Utilities Policy** 17, p. 267-275, 2009.

COSTA, M. B. F. Técnica derivada da teoria da resposta ao item aplicada ao setor de serviços. **Dissertação de Mestrado** - PPGMUE/UFPR. 2001.

DAS, J.; HAMMER, J. Which doctor? Combining vignettes and item response to measure clinical competence. **Journal of Development Economics** 78, p. 348-383, 2005.

DATTAKUMAR, R.; JAGADEESH, R. A review of literature on benchmarking, **Benchmarking: An International Journal**, v. 10, n. 3, p. 176-209, 2003.

DEROS, B. M.; YUSOF, S. M.; SALLEH, A. M. A benchmarking implementation framework for automotive manufacturing SMEs. **Benchmarking: An International Journal**, v. 13, n. 4, p. 396-430, 2006.

DONNELLY, M. A radical scoring system for the European Foundation for Quality Management Business Excellence Model. **Managerial Auditing Journal** 15/1/2, p. 8-11, 2000.

EMBRETSON, S. E.; REISE, S. P. **Item response theory for psychologists.** New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. 2000.

FONG, S. W.; CHENG, E. W. L.; HO, D. C. K. Benchmarking: a general reading for management practitioners. **Management Decision**, v. 36, Iss: 6, p. 407-418, 1998.

GARVIN, D. A. Building a learning organization. **Harvard Business Review**, v. 71, n. 4, p. 78-91, 1993.

GIL, A. C. **Como Elaborar Projetos de Pesquisa.** São Paulo: Atlas, 2002.

GRUPP, H. Indicators for national science and technology policy: how robust are composite indicators? **Research Policy** 33, p. 1373-1384, 2004.

GUTTMAN, L. The basis for scalogram analysis in STOUFFER, S. A.; GUTTMAN, L., SUCHMAN, E. A.; LAZARSELD, P. F.; STAR, S. A.; CLAUSEN, J. A. **Measurement and prediction** (Studies in social psychology in World War II, v. 4). Princeton NJ: Princeton University Press, 1950.

HABIBOV, N. N. Comparing and contrasting poverty reduction performance of social welfare programs across jurisdictions in Canada using Data Envelopment Analysis (DEA): An exploratory study of the era of devolution. **Evaluation and Program Planning** 33, p. 457-467, 2010.

HAMDAN, A. Evaluating the efficiency of 3PL logistics operations. **Int. J. Production Economics** 113, p. 235-244, 2008.

HO, W. Multiple criteria decision-making techniques in higher education. **International Journal of Educational Management** v. 20, n. 5, p. 319-337, 2006.

KAKATI, M. Strategic evaluation of advanced manufacturing technology. **Int. J. Production Economics** 53, p. 141-156, 1997.

KLEIN, R. Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB). **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 11, n. 40, p. 283-296, 2003.

KOTLER, P.; KELLER, K. L. **Marketing Management**. Prentice Hall Brasil, 12th Edition, 2006.

KYRÖ, P. Revising the concept and forms of benchmarking. **Benchmarking: An International Journal**, v. 10, n. 3, p. 210-225, 2003.

LEONG, W. H. Measuring the technical efficiency of banks in Singapore for the period 1993-99: an application and extension of the Bauer et al. (1997) technique. **ASEAN Economic Bulletin** v. 20, n. 3, p. 195-210, 2003.

LIKERT, R. A technique for the measurement of attitudes. **Archs Psychol.**, v. 22, n. 140, 1932.

LORD, F.M. **Applications of item response theory to practical testing problems**. Hillsdale: Lawrence Erlbaum Associates Inc., 1980.

- MASTERS, G. N. A Rasch model for partial credit scoring. **Psychometrika**, 47, p. 149-174, 1982.
- MCNAIR, C. J.; LEIBFRIED, K. H. J. **Benchmarking, A Tool for Continuous Improvement**. New York: Harper Business Press, 1992.
- MEHREGAN, M. R. An optimisational model of benchmarking. **Benchmarking: An International Journal** v. 17, n. 6, p. 876-888, 2010.
- MESBAH, M.; COLE, B. F.; LEE, M. L. T. **Statistical methods for quality of life studies: design, measurements and analysis**. Boston: Kluwer Academic Publishers, 2002.
- MISLEVY, R. J.; STOCKING, M. L. A Consumer's Guide to LOGISTIC and BILOG. **Applied Psychological Measurement**, 13 p. 57-75, 1989.
- MORIARTY, J. P. A theory of benchmarking. **Benchmarking: An International Journal**, v. 18, n. 4, p. 588-612, 2011.
- MURAKI, E. A generalized partial credit model : Application of an EM algorithm. **Applied Psychological Measurement**, 16, p. 159-176, 1992.
- NAOR, M. The globalization of operations in Eastern and Western countries: Unpacking the relationship between national and organizational culture and its impact on manufacturing performance. **Journal of Operations Management** 28, p. 194-205, 2010.
- NAYAR, P. Data Envelopment Analysis Comparison of Hospital Efficiency and Quality. **J Med Syst** 32, p. 193-199, 2008.
- OECD. **Programme for International Student Assessment (PISA)**. Disponível em: <http://www.oecd.org/pisa>. Acesso em: 02 janeiro 2012.
- OJANEN, V. Applying quality award criteria in R&D project assessment. **Int. J. Production Economics** 80, p. 119-128, 2002.
- PARTCHEV, I. **irtoys**: Simple interface to the estimation and plotting of IRT models. R package version 0.1.6. 2012. URL: <http://CRAN.R-project.org/package=irtoys>.
- PASIOURAS, F. Multicriteria classification models for the identification of targets and acquirers in the Asian banking sector. **European Journal of Operational Research** 204, p. 328-335, 2010.

PASQUALI, L. Princípios de elaboração de escalas psicológicas.

Revista de Psiquiatria Clínica, v. 25, n. 5, p. 206-221, 1998.

Disponível em:

<http://www.hcnet.usp.br/ipq/revista/vol25/n5/conc255a.htm>. Acesso em: 02 novembro 2012.

PASQUALI, L. **Psicometria - Teoria dos Testes na Psicologia e na Educação**. Petrópolis, RJ: Editora Vozes, 2003.

PORTER, M. E. **Competitive Advantage**. New York: Free Press, 1985.

PORTER, M. E. **How Competitive Forces Shape Strategy**, Harvard Business Review, March/April, 1979.

POTNIS, D. D. Measuring e-Governance as an innovation in the public sector. **Government Information Quarterly** 27, p. 41-48, 2010.

QFD Institute. The oficial source for QFD. Disponível em:

<http://www.qfdi.org/index.html>. Acesso em: 02 fevereiro 2013

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL: <http://www.R-project.org>.

SAMEJIMA, F. Estimation of latent ability using a response pattern of graded scores. **Psychometrika** Monograph Supplement, 1, 1969.

SAYEED, L. Can information technology improve managerial problem finding? **Information & Management** 27, p. 377-390, 1994.

SEVERINO, A. J. **Metodologia do Trabalho Científico**. São Paulo: Cortez, 2007.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis: UFSC, 2005.

Sistema Nacional de Avaliação da Educação Básica: SAEB 2001, Relatório Técnico. Consórcio Fundação Cesgranrio/Fundação Carlos Chagas, Rio de Janeiro, 2002.

SOHN, S. Y. Structural equation model for the evaluation of national funding on R&D project of SMEs in consideration with MBNQA criteria. **Evaluation and Program Planning** 30, p. 10-20, 2007.

SPENDOLINI, M. J. **The Benchmarking Book**. New York: AMACOM, 1992.

SPSS Inc. Released 2006. SPSS for Windows, Version 16.0. [Computer software]. Chicago, USA: SPSS Inc., 2006. URL: <http://www.ibm.com/software/analytics/spss/>.

TAN, H. B. Relative efficiency measures for the knowledge economies in the Asia Pacific region. **Journal of Modelling in Management** v.3, n. 2, p. 111-124, 2008.

TALLURI, S. Data Envelopment Analysis: Models and Extensions. 8 **Decision Line**, v. 31(3), p. 8-11, 2000. Disponível em: http://www.decisionsciences.org/decisionline/vol31/31_3/31_3pom.pdf. Acesso em: 02 fevereiro 2013.

TEZZA, R.; BORNIA, A. C.; ANDRADE, D. F. D. Measuring web usability using item response theory : Principles, features and opportunities. **Interacting with Computers**, v. 23, n. 2, p. 167-175, 2011.terra

THOMSON REUTERS. **EndNote X4**. [Computer software]. University of Lethbridge, 2010. URL: <http://www.endnote.com>.

VILLAS, M. V.; MACEDO-SOARES, T. D. L. van A. de; RUSSO, G. M. Bibliographical research method for business administration studies: a model based on scientific journal ranking. **BAR, Braz. Adm. Rev.**, Curitiba, v. 5, n. 2, 2008.

WARE, J. E.; BJORNER, J. B.; KOSINSKI, M. Practical Implications of Item Response Theory and Computerized Adaptive Testing. A Brief Summary of Ongoing Studies of Widely Used Headache Impact Scales. **Medical Care**, v. 38, 2000.

WATSON, G. H; YOUSSEF, M. A. A Perspective on Benchmarking: Gregory H. Watson in conversation with the Editor, **Benchmarking: An International Journal**, v. 1, n. 1, p. 5-10, 1994.

WATSON, G. H. **Strategic Benchmarking: How to Rate your Company's Performance against the World's Best**. New York: John Wiley & Sons Inc., 1993.

WILDE, P. E.; GERALD, J.; FRIEDMAN, D. R. Differential Response Patterns Affect Food-Security Prevalence Estimates for Households with and without Children. **J. Nutr.**, 134: p. 1910-1915. 2004.

WILSON, M. **Constructing Measures: An Item Response Modeling approach**, London: LEA, 2005.

YESAVAGE, J. A.; BRINK, T. L.; ROSE, T. L. Development and validation of a geriatric depression screening scale: a preliminary report. **J. Psychiat Res**, 17: p. 37-49, 1983.

ZIMOWSKI, M. F.; MURAKI, E.; MISLEVY, R. J.; BOCK, R. D. **BILOG-MG 3 for Windows: Multiple-group IRT analysis and test maintenance for binary items** [Computer software]. Skokie, IL: Scientific Software International, Inc., 2003. URL: <http://www.ssicentral.com>.

ANEXOS

A.1 INSTRUMENTO DE AVALIAÇÃO DE CURSOS DE GRADUAÇÃO PRESENCIAL E A DISTÂNCIA - PARCIAL

INSTRUÇÕES PARA PREENCHIMENTO

Este Instrumento subsidia os atos autorizativos de cursos – autorização, reconhecimento e renovação de reconhecimento – nos graus de tecnólogo, de licenciatura e de bacharelado para a modalidade presencial e a distância. De acordo com o art. 1º da Portaria Normativa 40/2007, consolidada em 29 de dezembro de 2010, a aplicação dos indicadores desse instrumento dar-se-á exclusivamente em meio eletrônico, no sistema e-MEC. Os avaliadores deverão considerar as orientações a seguir:

1. Atribuir conceitos de 1 a 5, em ordem crescente de excelência, a cada um dos indicadores de cada uma das três dimensões;
2. Considerar os critérios de análise dos respectivos indicadores da dimensão. A atribuição dos conceitos deve ser feita da forma seguinte:
Conceito - Descrição
 - 1 - Quando os indicadores da dimensão avaliada configuram um conceito NÃO EXISTENTE.
 - 2 - Quando os indicadores da dimensão avaliada configuram um conceito INSUFICIENTE.
 - 3 - Quando os indicadores da dimensão avaliada configuram um conceito SUFICIENTE.
 - 4 - Quando os indicadores da dimensão avaliada configuram um conceito MUITO BOM/MUITO BEM.
 - 5 - Quando os indicadores da dimensão avaliada configuram um conceito EXCELENTE.
3. Atribuir os conceitos a cada um dos indicadores. Os conceitos deverão ser contextualizados, com base nos indicadores, descritos de forma abrangente e coerentes no quadro “CONSIDERAÇÕES SOBRE A DIMENSÃO”;
4. Manter sempre a coerência entre o conceito atribuído aos indicadores e ao gerado na dimensão com a análise qualitativa;
5. Consultar o glossário sempre que necessário;

6. A contextualização da IES e do curso e a síntese preliminar devem conter, obrigatoriamente, os dados abaixo:

6.1. CONTEXTUALIZAÇÃO DA IES

- a) Nome da mantenedora;
- b) Base legal da mantenedora (endereço, razão social, registro no cartório e atos legais);
- c) Nome da IES;
- d) Base legal da IES (endereço, atos legais e data da publicação no DOU);
- e) Perfil e missão da IES;
- f) Dados socioeconômicos da região;
- g) Breve histórico da IES (criação, trajetória, áreas oferecidas no âmbito da graduação e da pós-graduação, áreas de atuação na extensão e áreas de pesquisa, se for o caso).

6.2. CONTEXTUALIZAÇÃO DO CURSO

- a) Nome do curso;
- b) Nome da mantida;
- c) Endereço de funcionamento do curso;
- d) Atos legais de Autorização, Reconhecimento e Renovação de Reconhecimento do curso, quando existirem;
- e) Número de vagas pretendidas ou autorizadas;
- f) Conceito Preliminar de Curso – CPC – e Conceito de Curso – CC –, quando houver;
- g) Turnos de funcionamento do curso (matutino, vespertino, noturno e integral);
- h) Carga horária total do curso (em horas e em hora/aula);
- i) Tempo mínimo e máximo para integralização;
- j) Identificação do (a) coordenador (a) do curso;
- k) Perfil do (a) coordenador (a) do curso (formação acadêmica, titulação, tempo de exercício na IES e na função de coordenador do curso);
- l) Composição, titulação, regime de trabalho e permanência sem interrupção dos integrantes do Núcleo Docente Estruturante – NDE;
- m) Tempo médio de permanência do corpo docente no curso (exceto para autorização). Somar o tempo de exercício no curso de todos os docentes e dividir pelo número total de docentes no curso, incluindo o tempo do (a) coordenador (a) do curso.

6.3. SÍNTESE PRELIMINAR

- a) Identificar a modalidade do curso;

- b) Realçar se há divergência no endereço de visita com o endereço do escritório de designação;
- c) Explicitar os documentos que serviram de base para análise da avaliação (PDI, PPC, relatórios de autoavaliação e demais relatórios da IES), e se estão dentro do prazo de validade;
- d) Observar as diligências e seu cumprimento;
- e) Verificar e comentar se o (a) coordenador (a) do curso apresentou justificativa procedente, ou não, ao CPC insatisfatório e se há coerência entre a justificativa apresentada e as ações propostas para sanear as possíveis deficiências (somente para Renovação de Reconhecimento de curso).

INFORMAÇÕES

1. O Conceito do Curso (CC) é calculado, pelo sistema e-MEC, com base na média aritmética ponderada dos conceitos das dimensões, os quais são resultados da média aritmética simples dos indicadores das respectivas dimensões.
2. Este instrumento possui indicadores com recurso de NSA, ou seja, “Não Se Aplica”. Quando o indicador não se aplicar à avaliação, a comissão deverá optar por NSA. Assim, este indicador não será considerado no cálculo da dimensão.
3. O termo Não Se Aplica – NSA, constante nos indicadores específicos, deverá ser analisado de acordo com as diretrizes curriculares do curso e será justificado pelo avaliador após análise do Projeto Pedagógico do Curso – PPC, do Plano de Desenvolvimento Institucional – PDI e das Diretrizes Curriculares Nacionais - DCNs.
4. Cada indicador apresenta, predominantemente, um objeto de análise.
5. Para os indicadores que contemplam o termo “Análise Sistêmica e Global”, a comissão deverá seguir somente os aspectos estabelecidos no respectivo indicador, baseados nas informações contidas no Projeto Pedagógico do Curso – PPC, no Plano de Desenvolvimento Institucional – PDI e nas e das Diretrizes Curriculares Nacionais – DCNs.

Dimensão 2: CORPO DOCENTE E TUTORIAL

Fontes de Consulta: Projeto Pedagógico do Curso, Formulário Eletrônico preenchido pela IES no e-MEC e Documentação Comprobatória.

Indicador	Conceito	Critério de Análise
2.1. Atuação do Núcleo Docente Estruturante – NDE	1	Quando o NDE não está previsto/implantado.
	2	Quando a atuação do NDE previsto/implantado é insuficiente considerando, em uma análise sistêmica e global, os aspectos: concepção, acompanhamento, consolidação e avaliação do PPC.
	3	Quando a atuação do NDE previsto/implantado é suficiente considerando, em uma análise sistêmica e global, os aspectos: concepção, acompanhamento, consolidação e avaliação do PPC.
	4	Quando a atuação do NDE previsto/implantado é muito boa considerando, em uma análise sistêmica e global, os aspectos: concepção, acompanhamento, consolidação e avaliação do PPC.
	5	Quando a atuação do NDE previsto/implantado é excelente considerando, em uma análise sistêmica e global, os aspectos: concepção, acompanhamento, consolidação e avaliação do PPC.

Indicador	Conceito	Critério de Análise
2.2. Atuação do (a) coordenador (a)	1	Quando não há atuação do (a) coordenador (a) do curso.
	2	Quando a atuação do (a) coordenador (a) é insuficiente considerando, em uma análise sistêmica e global, os aspectos: gestão do curso, relação com os docentes e discentes e representatividade nos colegiados superiores.
	3	Quando a atuação do (a) coordenador (a) é suficiente considerando, em uma análise sistêmica e global, os aspectos: gestão do curso, relação com os docentes e discentes e representatividade nos colegiados superiores.
	4	Quando a atuação do (a) coordenador (a) é muito boa considerando, em uma análise sistêmica e global, os aspectos: gestão do curso, relação com os docentes e discentes e representatividade nos colegiados superiores.
	5	Quando a atuação do (a) coordenador (a) é excelente considerando, em uma análise sistêmica e global, os aspectos: gestão do curso, relação com os docentes e discentes e representatividade nos colegiados superiores.

Indicador	Conceito	Critério de Análise
2.3. Experiência do (a) coordenador (a) do curso em cursos a distância (Indicador específico para cursos a distância)	1	Quando o (a) coordenador (a) possui menos de 1 ano de experiência em cursos a distância.
	2	Quando o (a) coordenador (a) possui experiência em cursos a distância maior ou igual a 1 ano e menor que 2 anos.
	3	Quando o (a) coordenador (a) possui experiência em cursos a distância maior ou igual a 2 anos e menor que 3 anos.
	4	Quando o (a) coordenador (a) possui experiência em cursos a distância maior ou igual a 3 anos e menor que 4 anos.
	5	Quando o (a) coordenador (a) possui experiência em cursos a distância maior ou igual a 4 anos.

Indicador	Conceito	Critério de Análise
2.4. Experiência profissional, de magistério superior e de gestão acadêmica do (a) coordenador (a)	1	Quando o (a) coordenador (a) possui experiência profissional, de magistério superior e de gestão acadêmica, somadas, menor que 1 ano ou menos de 1 ano de magistério superior .
	2	Quando o (a) coordenador (a) possui experiência profissional, de magistério superior e de gestão acadêmica, somadas, maior ou igual a 1 ano e menor que 4 anos sendo, no mínimo, 1 ano de magistério superior .
	3	Quando o (a) coordenador (a) possui experiência profissional, de magistério superior e de gestão acadêmica, somadas, maior ou igual a 4 anos e menor que 7 anos sendo, no mínimo, 1 ano de magistério superior .
	4	Quando o (a) coordenador (a) possui experiência profissional, de magistério superior e de gestão acadêmica, somadas, maior ou igual a 7 anos e menor que 10 anos sendo, no mínimo, 1 ano de magistério superior .
	5	Quando o (a) coordenador (a) possui experiência profissional, de magistério superior e de gestão acadêmica, somadas, maior ou igual a 10 anos sendo, no mínimo, 1 ano de magistério superior .

Indicador	Conceito	Critério de Análise
2.5. Regime de trabalho do (a) coordenador (a) do curso NSA para cursos a distância, obrigatório para cursos presenciais	1	Quando o regime de trabalho previsto/implantado do (a) coordenador (a) não é de tempo parcial ou integral , ou a relação entre o número de vagas anuais pretendidas/autorizadas e as horas semanais dedicadas à coordenação é maior que 25 , ou não é respeitado o limite mínimo de 10 horas semanais dedicadas à coordenação do curso.
	2	Quando o regime de trabalho previsto/implantado do (a) coordenador (a) é de tempo parcial ou integral , sendo que a relação entre o número de vagas anuais pretendidas/autorizadas e as horas semanais dedicadas à coordenação é maior que 20 e menor ou igual a 25 .
	3	Quando o regime de trabalho previsto/implantado do (a) coordenador (a) é de tempo parcial ou integral , sendo que a relação entre o número de vagas anuais pretendidas/autorizadas e as horas semanais dedicadas à coordenação é maior que 15 e menor ou igual a 20 .
	4	Quando o regime de trabalho previsto/implantado do (a) coordenador (a) é de tempo parcial ou integral , sendo que a relação entre o número de vagas anuais pretendidas/autorizadas e as horas semanais dedicadas à coordenação é maior que 10 e menor ou igual a 15 .
	5	Quando o regime de trabalho previsto/implantado do (a) coordenador (a) é de tempo parcial ou integral , sendo que a relação entre o número de vagas anuais pretendidas/autorizadas e as horas semanais dedicadas à coordenação é menor ou igual a 10 .

Indicador	Conceito	Critério de Análise
2.6. Carga horária de coordenação de curso NSA para cursos presenciais, obrigatório para cursos a distância	1	Quando a carga horária prevista/implantada para o (a) coordenador (a) for menor que 10 horas .
	2	Quando a carga horária prevista/implantada para o (a) coordenador (a) do curso for maior ou igual a 10 e menor que 15 horas semanais dedicadas totalmente à coordenação.
	3	Quando a carga horária prevista/implantada para o (a) coordenador (a) do curso for maior ou igual a 15 e menor que 20 horas semanais dedicadas totalmente à coordenação.
	4	Quando a carga horária prevista/implantada para o (a) coordenador (a) do curso for maior ou igual a 20 e menor que 25 horas semanais dedicadas totalmente à coordenação.
	5	Quando a carga horária prevista/implantada para o (a) coordenador (a) do curso for maior ou igual a 25 horas semanais dedicadas totalmente à coordenação.

Indicador	Conceito	Critério de Análise
2.7. Titulação do corpo docente do curso (Para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas)	1	Quando o percentual dos docentes do curso com titulação obtida em programas de pós-graduação <i>stricto sensu</i> é menor que 15% .
	2	Quando o percentual dos docentes do curso com titulação obtida em programas de pós-graduação <i>stricto sensu</i> é maior ou igual a 15% e menor que 30% .
	3	Quando o percentual dos docentes do curso com titulação obtida em programas de pós-graduação <i>stricto sensu</i> é maior ou igual a 30% e menor que 50% .
	4	Quando o percentual dos docentes do curso com titulação obtida em programas de pós-graduação <i>stricto sensu</i> é maior ou igual a 50% e menor que 75% .
	5	Quando o percentual dos docentes do curso com titulação obtida em programas de pós-graduação <i>stricto sensu</i> é maior ou igual a 75% .

Indicador	Conceito	Critério de Análise
2.8. Titulação do corpo docente do curso – percentual de doutores (Para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas)	1	Quando não há doutores do curso.
	2	Quando o percentual de doutores do curso é menor ou igual a 10% .
	3	Quando o percentual de doutores do curso é maior que 10% e menor ou igual a 20% .
	4	Quando o percentual de doutores do curso é maior que 20% e menor ou igual a 35% .
	5	Quando o percentual de doutores do curso é maior que 35% .

Indicador	Conceito	Critério de Análise
2.9. Regime de trabalho do corpo docente do curso (Para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas) (Para os cursos de Medicina, os critérios de análise passam a figurar da seguinte maneira: Conceito 1 – menor que 50% Conceito 2 – maior ou igual a 50% e menor que 60% Conceito 3 – maior ou igual a 60% e menor que 70% Conceito 4 – maior ou igual a 70% e menor que 80% Conceito 5 – maior ou igual a 80%)	1	Quando o percentual do corpo docente previsto/efetivo com regime de trabalho de tempo parcial ou integral é menor que 20% .
	2	Quando o percentual do corpo docente previsto/efetivo com regime de trabalho de tempo parcial ou integral é maior ou igual a 20% e menor que 33% .
	3	Quando o percentual do corpo docente previsto/efetivo com regime de trabalho de tempo parcial ou integral é maior ou igual a 33% e menor que 60% .
	4	Quando o percentual do corpo docente previsto/efetivo com regime de trabalho de tempo parcial ou integral é maior ou igual a 60% e menor que 80% .
	5	Quando o percentual do corpo docente previsto/efetivo com regime de trabalho de tempo parcial ou integral é maior ou igual que 80% .

Indicador	Conceito	Critério de Análise
<p>2.10. Experiência profissional do corpo docente</p> <p>(Para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas)</p> <p>NSA para egressos de cursos de licenciatura</p> <p>(Para os cursos de Medicina, os critérios de análise passam a figurar da seguinte maneira: Conceito 1 – menor que 40% possui, pelo menos, 5 anos Conceito 2 – maior ou igual a 40% e menor que 50% possui, pelo menos, 5 anos Conceito 3 – maior ou igual a 50% e menor que 60% possui, pelo menos, 5 anos Conceito 4 – maior ou igual a 60% e menor que 70% possui, pelo menos, 5 anos Conceito 5 – maior ou igual a 70% possui, pelo menos, 5 anos)</p>	1	Quando um contingente menor que 20% do corpo docente previsto/efetivo possui experiência profissional (excluída as atividades no magistério superior) de, pelo menos, 2 anos para bacharelados/licenciaturas ou 3 anos para cursos superiores de tecnologia.
	2	Quando um contingente maior ou igual a 20% e menor que 40% do corpo docente previsto/efetivo possui experiência profissional (excluída as atividades no magistério superior) de, pelo menos, 2 anos para bacharelados/licenciaturas ou 3 anos para cursos superiores de tecnologia.
	3	Quando um contingente maior ou igual a 40% e menor que 60% do corpo docente previsto/efetivo possui experiência profissional (excluída as atividades no magistério superior) de, pelo menos, 2 anos para bacharelados/licenciaturas ou 3 anos para cursos superiores de tecnologia.
	4	Quando um contingente maior ou igual a 60% e menor que 80% do corpo docente previsto/efetivo possui experiência profissional (excluída as atividades no magistério superior) de, pelo menos, 2 anos para bacharelados/licenciaturas ou 3 anos para cursos superiores de tecnologia.
	5	Quando um contingente maior ou igual a 80% do corpo docente previsto/efetivo possui experiência profissional (excluída as atividades no magistério superior) de, pelo menos, 2 anos para bacharelados/licenciaturas ou 3 anos para cursos superiores de tecnologia.

Indicador	Conceito	Critério de Análise
<p>2.11. Experiência no exercício da docência na educação básica</p> <p>(Para fins de autorização, considerar os docentes previstos para os dois primeiros anos do curso)</p> <p>Obrigatório para cursos de licenciatura, NSA para os demais</p>	1	Quando um contingente menor que 20% do corpo docente previsto/efetivo tem, pelo menos, 3 anos de experiência no exercício da docência na educação básica.
	2	Quando um contingente maior ou igual a 20% e menor que 30% do corpo docente previsto/efetivo tem, pelo menos, 3 anos de experiência no exercício da docência na educação básica.
	3	Quando um contingente maior ou igual a 30% e menor que 40% do corpo docente previsto/efetivo tem, pelo menos, 3 anos de experiência no exercício da docência na educação básica.
	4	Quando um contingente maior ou igual a 40% e menor que 50% do corpo docente previsto/efetivo tem, pelo menos, 3 anos de experiência no exercício da docência na educação básica.
	5	Quando um contingente maior ou igual a 50% do corpo docente previsto/efetivo tem, pelo menos, 3 anos de experiência no exercício da docência na educação básica.

Indicador	Conceito	Critério de Análise
2.12. Experiência de magistério superior do corpo docente (Para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas) (Para os cursos de Medicina, os critérios de análise passam a figurar da seguinte maneira: Conceito 1 – menor que 40% possui, pelo menos, 5 anos Conceito 2 – maior ou igual a 40% e menor que 50% possui, pelo menos, 5 anos Conceito 3 – maior ou igual a 50% e menor que 60% possui, pelo menos, 5 anos Conceito 4 – maior ou igual a 60% e menor que 70% possui, pelo menos, 5 anos Conceito 5 – maior ou igual a 70% possui, pelo menos, 5 anos)	1	Quando um contingente menor que 20% do corpo docente previsto/efetivo possui experiência de magistério superior de, pelo menos, 3 anos para bacharelados/licenciaturas ou 2 anos para cursos superiores de tecnologia.
	2	Quando um contingente maior ou igual a 20% e menor que 40% do corpo docente previsto/efetivo possui experiência de magistério superior de, pelo menos, 3 anos para bacharelados/licenciaturas ou 2 anos para cursos superiores de tecnologia.
	3	Quando um contingente maior ou igual a 40% e menor que 60% do corpo docente previsto/efetivo possui experiência de magistério superior de, pelo menos, 3 anos para bacharelados/licenciaturas ou 2 anos para cursos superiores de tecnologia.
	4	Quando um contingente maior ou igual a 60% e menor que 80% do corpo docente previsto/efetivo possui experiência de magistério superior de, pelo menos, 3 anos para bacharelados/licenciaturas ou 2 anos para cursos superiores de tecnologia.
	5	Quando um contingente maior ou igual a 80% do corpo docente previsto/efetivo possui experiência de magistério superior de, pelo menos, 3 anos para bacharelados/licenciaturas ou 2 anos para cursos superiores de tecnologia.

Indicador	Conceito	Critério de Análise
2.13. Relação entre o número de docentes e o número de estudantes NSA para cursos presenciais. Obrigatório para cursos a distância (relação entre o número de docentes e o número de estudantes equivalente 40h em dedicação à EAD)	1	Quando a média entre o número de docentes do curso (equivalentes 40h) e o número de vagas previstas/implantadas é de 1 docente para 161 vagas ou mais .
	2	Quando a média entre o número de docentes do curso (equivalentes 40h) e o número de vagas previstas/implantadas é de 1 docente para 151 a 160 vagas .
	3	Quando a média entre o número de docentes do curso (equivalentes 40h) e o número de vagas previstas/implantadas é de 1 docente para 141 a 150 vagas .
	4	Quando a média entre o número de docentes do curso (equivalentes 40h) e o número de vagas previstas/implantadas é de 1 docente para 131 a 140 vagas .
	5	Quando a média entre o número de docentes do curso (equivalentes 40h) e o número de vagas previstas/implantadas é de 1 docente para 130 .

Indicador	Conceito	Critério de Análise
2.14. Funcionamento do colegiado de curso ou equivalente	1	Quando o colegiado não está previsto/implantado.
	2	Quando o funcionamento do colegiado previsto/implantado está regulamentado/institucionalizado, de maneira insuficiente , considerando, em uma análise sistêmica e global, os aspectos: representatividade dos segmentos, periodicidade das reuniões, registros e encaminhamento das decisões.
	3	Quando o funcionamento do colegiado previsto/implantado está regulamentado/institucionalizado, de maneira suficiente , considerando, em uma análise sistêmica e global, os aspectos: representatividade dos segmentos, periodicidade das reuniões, registros e encaminhamento das decisões.
	4	Quando o funcionamento do colegiado previsto/implantado está muito bem regulamentado/institucionalizado considerando, em uma análise sistêmica e global, os aspectos: representatividade dos segmentos, periodicidade das reuniões, registros e encaminhamento das decisões.
	5	Quando o funcionamento do colegiado previsto/implantado está regulamentado/institucionalizado, de maneira excelente , considerando, em uma análise sistêmica e global, os aspectos: representatividade dos segmentos, periodicidade das reuniões, registros e encaminhamento das decisões.

Indicador	Conceito	Critério
2.15. Produção científica, cultural, artística ou tecnológica (para fins de autorização, considerar os docentes previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas)	1	Quando menos de 50% dos docentes não têm produção nos últimos 3 anos.
	2	Quando pelo menos 50% dos docentes têm entre 1 a 3 produções nos últimos 3 anos.
	3	Quando pelo menos 50% dos docentes têm entre 4 a 6 produções nos últimos 3 anos.
	4	Quando pelo menos 50% dos docentes têm entre 7 a 9 produções nos últimos 3 anos.
	5	Quando pelo menos 50% dos docentes têm mais de 9 produções nos últimos 3 anos.

Indicador	Conceito	Critério de Análise
2.16. Titulação e formação do corpo de tutores do curso (Para fins de autorização, considerar os tutores previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas) NSA para cursos presenciais. Obrigatório para cursos a distância e presenciais, reconhecidos, que ofertam até 20% da carga horária total do curso na modalidade a distância, conforme Portaria 4.059/2004	1	Quando existem tutores não graduados .
	2	Quando todos os tutores previstos/efetivos são graduados .
	3	Quando todos os tutores previstos/efetivos são graduados na área .
	4	Quando todos os tutores previstos/efetivos são graduados na área , sendo que, no mínimo, 30% têm titulação obtida em programas de pós-graduação lato sensu .
	5	Quando todos os tutores previstos/efetivos são graduados na área , sendo que, no mínimo, 30% têm titulação obtida em programas de pós-graduação stricto sensu .

Indicador	Conceito	Critério de Análise
2.17. Experiência do corpo de tutores em educação a distância (Para fins de autorização, considerar os tutores previstos para o primeiro ano do curso, se CSTs, ou dois primeiros anos, se bacharelados/licenciaturas) NSA para cursos presenciais. Obrigatório para cursos a distância e presenciais, reconhecidos, que ofertam até 20% da carga horária total do curso na modalidade a distância, conforme Portaria 4.059/2004	1	Quando o percentual de tutores do curso previstos/efetivos que possui experiência mínima de 3 anos em cursos a distância é menor que 40% .
	2	Quando o percentual de tutores do curso previstos/efetivos que possui experiência mínima de 3 anos em cursos a distância é maior ou igual a 40% e menor que 50% .
	3	Quando o percentual de tutores do curso previstos/efetivos que possui experiência mínima de 3 anos em cursos a distância é maior ou igual a 50% e menor que 60% .
	4	Quando o percentual de tutores do curso previstos/efetivos que possui experiência mínima de 3 anos em cursos a distância é maior ou igual a 60% e menor que 70% .
	5	Quando o percentual de tutores do curso previstos/efetivos que possui experiência mínima de 3 anos em cursos a distância é maior ou igual a 70% .

Indicador	Conceito	Critério de Análise
2.18. Relação docentes e tutores – presenciais e a distância – por estudante NSA para cursos presenciais. Obrigatório para cursos a distância e presenciais, reconhecidos, que ofertam até 20% da carga horária total do curso na modalidade a distância, conforme Portaria 4.059 /2004	1	Quando a relação entre o número de estudantes e o total de docentes mais tutores (presenciais e a distância) previstos/contratados é maior que 60 .
	2	Quando a relação entre o número de estudantes e o total de docentes mais tutores (presenciais e a distância) previstos/contratados é maior que 50 e menor ou igual a 60 .
	3	Quando a relação entre o número de estudantes e o total de docentes mais tutores (presenciais e a distância) previstos/contratados é maior que 40 e menor ou igual a 50 .
	4	Quando a relação entre o número de estudantes e o total de docentes mais tutores (presenciais e a distância) previstos/contratados é maior que 30 e menor ou igual a 40 .
	5	Quando a relação entre o número de estudantes e o total de docentes mais tutores (presenciais e a distância) previstos/contratados é menor ou igual a 30 .

Indicador	Conceito	Critério de Análise
2.19. Responsabilidade docente pela supervisão da assistência médica Obrigatório para o curso de Medicina, NSA para os demais cursos	1	Quando a porcentagem dos docentes responsáveis pelas atividades de ensino envolvendo pacientes que se responsabiliza pela supervisão da assistência médica a elas vinculadas é menor que 25% .
	2	Quando a porcentagem dos docentes responsáveis pelas atividades de ensino envolvendo pacientes que se responsabiliza pela supervisão da assistência médica a elas vinculadas é maior ou igual a 25% e menor que 50% , sendo que, destes, pelo menos 30% dos docentes supervisionam e são responsáveis pelos serviços clínico-cirúrgicos frequentados pelos alunos.
	3	Quando a porcentagem dos docentes responsáveis pelas atividades de ensino envolvendo pacientes que se responsabiliza pela supervisão da assistência médica a elas vinculadas é maior ou igual a 50% e menor que 75% , sendo que, destes, pelo menos 30% dos docentes supervisionam e são responsáveis pelos serviços clínico-cirúrgicos frequentados pelos alunos.
	4	Quando a porcentagem dos docentes responsáveis pelas atividades de ensino envolvendo pacientes que se responsabiliza pela supervisão da assistência médica a elas vinculadas é maior ou igual a 75% e menor que 90% , sendo que, destes, pelo menos 30% dos docentes supervisionam e são responsáveis pelos serviços clínico-cirúrgicos frequentados pelos alunos.
	5	Quando a porcentagem dos docentes responsáveis pelas atividades de ensino envolvendo pacientes que se responsabiliza pela supervisão da assistência médica a elas vinculadas é maior ou igual a 90% , sendo que, destes, pelo menos 30% dos docentes supervisionam e são responsáveis pelos serviços clínico-cirúrgicos frequentados pelos alunos.

Indicador	Conceito	Critério de Análise
2.20. Núcleo de apoio pedagógico e experiência docente Obrigatório para o curso de Medicina, NSA para os demais cursos	1	Quando o núcleo de apoio pedagógico e experiência docente não está previsto/implantado, ou quando menos de 6 das áreas do curso estão representadas no núcleo existente.
	2	Quando o núcleo de apoio pedagógico e experiência docente previsto/implantado é composto por docentes do curso com, no mínimo, 2 anos de experiência docente, ou não cobre todas as áreas temáticas do curso.
	3	Quando o núcleo de apoio pedagógico e experiência docente previsto/implantado é composto por docentes do curso com, no mínimo, 3 anos de experiência docente, cobrindo todas as áreas temáticas do curso.
	4	Quando o núcleo de apoio pedagógico e experiência docente previsto/implantado é composto por docentes do curso com, no mínimo, 4 anos de experiência docente, cobrindo todas as áreas temáticas do curso.
	5	Quando o núcleo de apoio pedagógico e experiência docente previsto/implantado é composto por docentes do curso com, no mínimo, 5 anos de experiência docente, cobrindo todas as áreas temáticas do curso.

A.2 SCRIPTS DE LEITURA DOS MICRODADOS - IES

```

*****
* MEC/INEP
* DIRETORIA DE ESTATÍSTICAS EDUCACIONAIS
* Coordenação de Indicadores e Controle de Qualidade da
* Educação Superior
*-----*
* PROGRAMA PARA LEITURA DOS MICRODADOS DO CENSO DA EDUCAÇÃO
* SUPERIOR
* IES
*-----*
* DESCRIÇÃO:
* Microdados da Base de Dados do Censo da Educação Superior
* 2010
*****
* Obs:
* 1)Para execução desse programa é recomendável que o arquivo
* IES.txt (CD-ROM pasta DADOS) seja descompactado
* no diretório(C:\ )do computador e este programa seja salvo
* no mesmo.
* Ao terminar esses procedimentos execute o programa salvo
* utilizando as variáveis de interesse.
*****

```

.

* Leitura das variáveis.

```

GET DATA TYPE = TXT
  /FILE = 'C:\IES.TXT'
  /FIXCASE = 1
  /ARRANGEMENT = FIXED
  /FIRSTCASE = 1
  /IMPORTCASE = ALL
  /VARIABLES =
  /1
CO_IES 0-7 F8
CO_MANTENEDORA 8-15 F8
CO_CATEGORIA_ADMINISTRATIVA 16-23 F8
DS_CATEGORIA_ADMINISTRATIVA 24-123 A100
CO_ORGANIZACAO_ACADEMICA 124-131 F8
DS_ORGANIZACAO_ACADEMICA 132-231 A100
CO_MUNICIPIO_IES 232-239 F8
NO_MUNICIPIO_IES 240-389 A150
CO_UF 390-397 F8
SGL_UF 398-399 A2
NO_REGIAO 400-429 A30
IN_CAPITAL 430-437 F8
QT_TEC_TOTAL 438-445 F8
QT_TEC_FUND_INCOMP_FEM 446-453 F8
QT_TEC_FUND_INCOMP_MASC 454-461 F8
QT_TEC_FUND_COMP_FEM 462-469 F8
QT_TEC_FUND_COMP_MASC 470-477 F8
QT_TEC_MEDIO_FEM 478-485 F8

```

QT_TEC_MEDIO_MASC	486-493	F8
QT_TEC_SUPERIOR_FEM	494-501	F8
QT_TEC_SUPERIOR_MASC	502-509	F8
QT_TEC_ESPECIAL_FEM	510-517	F8
QT_TEC_ESPECIAL_MASC	518-525	F8
QT_TEC_MESTRADO_FEM	526-533	F8
QT_TEC_MESTRADO_MASC	534-541	F8
QT_TEC_DOUTORADO_FEM	542-549	F8
QT_TEC_DOUTORADO_MASC	550-557	F8
IN_ACESSO_PORTAL_CAPES	558-565	F8
IN_ACESSO_OUTRAS_BASES	566-573	F8
IN_REFERENTE	574-581	F8
VL_RECEITA_PROPRIA	582-595	COMMA14.2
VL_TRANSFERENCIA	596-609	COMMA14.2
VL_OUTRA_RECEITA	610-623	COMMA14.2
VL_DES_PESSOAL_REM_DOCENTE	624-637	COMMA14.2
VL_DES_PESSOAL_REM_TECNICO	638-651	COMMA14.2
VL_DES_PESSOAL_ENCARGO	652-665	COMMA14.2
VL_DES_CUSTEIO	666-679	COMMA14.2
VL_DES_INVESTIMENTO	680-693	COMMA14.2
VL_DES_PESQUISA	694-707	COMMA14.2
VL_DES_OUTRAS	708-721	COMMA14.2
.		
CACHE.		
EXECUTE.		

A.3 SCRIPTS DE LEITURA DOS MICRODADOS - CURSOS

```

*****
* MEC/INEP
* DIRETORIA DE ESTATÍSTICAS EDUCACIONAIS
* Coordenação de Indicadores e Controle de Qualidade da
* Educação Superior
* -----
* PROGRAMA PARA LEITURA DOS MICRODADOS DO CENSO DA EDUCAÇÃO
* SUPERIOR
* CURSO
* -----
* DESCRIÇÃO:
* Microdados da Base de Dados do Censo da Educação Superior
* 2010
*****
* Obs:
* 1)Para execução desse programa é recomendável que o arquivo
* CURSO.txt (CD-ROM pasta DADOS) seja descompactado
* no diretório(C:\ )do computador e este programa seja salvo
* no mesmo.
* Ao terminar esses procedimentos execute o programa salvo
* utilizando as variáveis de interesse.
*****

```

.

* Leitura das variáveis.

```

GET DATA TYPE = TXT
  /FILE = 'C:\CURSO.TXT'
  /FIXCASE = 1
  /ARRANGEMENT = FIXED
  /FIRSTCASE = 1
  /IMPORTCASE = ALL
  /VARIABLES =
  /1
CO_IES 0-7 F8
CO_CATEGORIA_ADMINISTRATIVA 8-15 F8
DS_CATEGORIA_ADMINISTRATIVA 16-115 A100
CO_ORGANIZACAO_ACADEMICA 116-123 F8
DS_ORGANIZACAO_ACADEMICA 124-223 A100
CO_MUNICIPIO 224-231 F8
CO_UF 232-239 F8
CO_CURSO 240-247 F8
NO_CURSO 248-447 A200
CO_OCDE 448-459 A12
NO_OCDE 460-579 A120
CO_OCDE_AREA_GERAL 580-591 A12
NO_AREA_GERAL 592-711 A120
CO_OCDE_AREA_ESPECIFICA 712-723 A12
NO_AREA_ESPECIFICA 724-843 A120
CO_OCDE_AREA_DETALHADA 844-855 A12
NO_AREA_DETALHADA 856-975 A120
CO_GRAU_ACADEMICO 976-983 F8

```

CO_MODALIDADE_ENSINO	984-991	F8
CO_NIVEL_ACADEMICO	992-999	F8
IN_GRATUITO	1000-1007	F8
TP_ATRIBUTO_INGRESSO	1008-1015	F8
CO_LOCAL_OFERTA_IES	1016-1023	F8
NU_CARGA_HORARIA	1024-1031	F8
DT_INICIO_FUNCIONAMENTO	1032-1041	A10
IN_AJUDA_DEFICIENTE	1042-1049	F8
IN_MATERIAL_DIGITAL	1050-1057	F8
IN_MATERIAL_IMPRESSO	1058-1065	F8
IN_MATERIAL_AUDIO	1066-1073	F8
IN_MATERIAL_BRILLE	1074-1081	F8
IN_DISCIPLINA_LIBRAS	1082-1089	F8
IN_GUIA_INTERPRETE	1090-1097	F8
IN_MATERIAL_LIBRAS	1098-1105	F8
IN_SINTESE_VOZ	1106-1113	F8
IN_TRADUTOR_LIBRAS	1114-1121	F8
IN_INTEGRAL_CURSO	1122-1129	F8
IN_MATUTINO_CURSO	1130-1137	F8
IN_NOTURNO_CURSO	1138-1145	F8
IN_VESPERTINO_CURSO	1146-1153	F8
NU_PERC_CARGA_HOR_DISTANCIA	1154-1161	F8
NU_INTEGRALIZACAO_MATUTINO	1162-1169	F8
NU_INTEGRALIZACAO_VESPERTINO	1170-1177	F8
NU_INTEGRALIZACAO_NOTURNO	1178-1185	F8
NU_INTEGRALIZACAO_INTEGRAL	1186-1193	F8
TP_MOTIVO_SEM_VINCULO	1194-1201	F8
CO_CURSO_REPRESENTADO	1202-1209	F8
IN_OFERECE_DISC_DISTANCIA	1210-1217	F8
IN_UTILIZA_LABORATORIO	1218-1225	F8
QT_INSCRITOS_ANO_EAD	1226-1233	F8
QT_VAGAS_ANUAL_EAD	1234-1241	F8
QT_VAGAS_INTEGRAL_PRES	1242-1249	F8
QT_VAGAS_MATUTINO_PRES	1250-1257	F8
QT_VAGAS_NOTURNO_PRES	1258-1265	F8
QT_VAGAS_VESPERTINO_PRES	1266-1273	F8
QT_INSCRITOS_MATUTINO_PRES	1274-1281	F8
QT_INSCRITOS_VESPERTINO_PRES	1282-1289	F8
QT_INSCRITOS_NOTURNO_PRES	1290-1297	F8
QT_INSCRITOS_INTEGRAL_PRES	1298-1305	F8
QT_MATRICULA_CURSO	1306-1313	F8
QT_CONCLUINTE_CURSO	1314-1321	F8
QT_INGRESSO_CURSO	1322-1329	F8
QT_INGRESSO_PROCESSO_SELETIVO	1330-1337	F8
QT_INGRESSO_OUTRA_FORMA	1338-1345	F8

.
 CACHE.
 EXECUTE.

A.4 SCRIPTS DE LEITURA DOS MICRODADOS - DOCENTES

```

*****
* MEC/INEP
* DIRETORIA DE ESTATÍSTICAS EDUCACIONAIS
* Coordenação de Indicadores e Controle de Qualidade da
* Educação Superior
* -----
* PROGRAMA PARA LEITURA DOS MICRODADOS DO CENSO DA EDUCAÇÃO
* SUPERIOR
* DOCENTE
* -----
* DESCRIÇÃO:
* Microdados da Base de Dados do Censo da Educação Superior
* 2010
*****
* Obs:
* 1)Para execução desse programa é recomendável que o arquivo
* DOCENTE.txt (CD-ROM pasta DADOS) seja descompactado
* no diretório(C:\ )do computador e este programa seja salvo
* no mesmo.
* Ao terminar esses procedimentos execute o programa salvo
* utilizando as variáveis de interesse.
*****

```

.

* Leitura das variáveis.

```

GET DATA TYPE = TXT
  /FILE = 'C:\DOCENTE.TXT'
  /FIXCASE = 1
  /ARRANGEMENT = FIXED
  /FIRSTCASE = 1
  /IMPORTCASE = ALL
  /VARIABLES =
  /1
CO_IES 0-7 F8
CO_CATEGORIA_ADMINISTRATIVA 8-15 F8
DS_CATEGORIA_ADMINISTRATIVA 16-115 A100
CO_ORGANIZACAO_ACADEMICA 116-123 F8
DS_ORGANIZACAO_ACADEMICA 124-223 A100
IN_CAPITAL 224-231 F8
CO_DOCENTE_IES 232-239 F8
CO_DOCENTE 240-253 A14
CO_SITUACAO_DOCENTE 254-261 F8
CO_ESCOLARIDADE_DOCENTE 262-269 F8
DS_ESCOLARIDADE_DOCENTE 270-283 A14
CO_REGIME_TRABALHO 284-291 F8
DS_REGIME_TRABALHO 292-329 A38
IN_SEXO_DOCENTE 330-337 F8
DS_SEXO_DOCENTE 338-346 A9
NU_ANO_DOCENTE_NASC 347-350 A4
NU_MES_DOCENTE_NASC 351-352 A2
NU_DIA_DOCENTE_NASC 353-354 A2

```

NU_IDADE_DOCENTE	355-362	F8
CO_COR_RACA_DOCENTE	363-370	F8
DS_COR_RACA_DOCENTE	371-394	A24
CO_PAIS_DOCENTE	395-402	F8
CO_NACIONALIDADE_DOCENTE	403-410	F8
CO_UF_NASCIMENTO	411-418	F8
CO_MUNICIPIO_NASCIMENTO	419-426	F8
IN_DOCENTE_DEFICIENCIA	427-434	F8
IN_CEGUEIRA	435-442	F8
IN_BAIXA_VISAO	443-450	F8
IN_SURDEZ	451-458	F8
IN_DEFICIENCIA_AUDITIVA	459-466	F8
IN_DEFICIENCIA_FISICA	467-474	F8
IN_SURDOCEGUEIRA	475-482	F8
IN_DEFICIENCIA_MULTIPLA	483-490	F8
IN_DEFICIENCIA_INTELECTUAL	491-498	F8
IN_ATU_EAD	499-506	F8
IN_ATU_EXTENSAO	507-514	F8
IN_ATU_GESTAO	515-522	F8
IN_ATU_GRAD_PRESENCIAL	523-530	F8
IN_ATU_POS_EAD	531-538	F8
IN_ATU_POS_PRESENCIAL	539-546	F8
IN_ATU_SEQUENCIAL	547-554	F8
IN_ATU_PESQUISA	555-562	F8
IN_BOLSA_PESQUISA	563-570	F8
IN_SUBSTITUTO	571-578	F8
IN_EXERCICIO_DT_REF	579-586	F8
IN_VISITANTE	587-594	F8
IN_VISITANTE_IFES_VINCULO	595-602	F8
DT_CADASTRO	603-612	A10

.

CACHE.
EXECUTE.

A.5 SCRIPTS DE LEITURA DOS MICRODADOS - ALUNOS

```

*****
* MEC/INEP
* DIRETORIA DE ESTATÍSTICAS EDUCACIONAIS
* Coordenação de Indicadores e Controle de Qualidade da
* Educação Superior
* -----
* PROGRAMA PARA LEITURA DOS MICRODADOS DO CENSO DA EDUCAÇÃO
* SUPERIOR
* ALUNOS
* -----
* DESCRIÇÃO:
* Microdados da Base de Dados do Censo da Educação Superior
* 2010
*****
* Obs:
* 1)Para execução desse programa é recomendável que o arquivo
* ALUNO.txt (CD-ROM pasta DADOS) seja descompactado
* no diretório(C:\ )do computador e este programa seja salvo
* no mesmo.
* Ao terminar esses procedimentos execute o programa salvo
* utilizando as variáveis de interesse.
*****

```

.

* Leitura das variáveis.

```

GET DATA TYPE = TXT
  /FILE = 'C:\ALUNO.TXT'
  /FIXCASE = 1
  /ARRANGEMENT = FIXED
  /FIRSTCASE = 1
  /IMPORTCASE = ALL
  /VARIABLES =
  /1
CO_IES 0-7 F8
CO_CATEGORIA_ADMINISTRATIVA 8-15 F8
DS_CATEGORIA_ADMINISTRATIVA 16-115 A100
CO_ORGANIZACAO_ACADEMICA 116-123 F8
DS_ORGANIZACAO_ACADEMICA 124-223 A100
CO_CURSO 224-231 F8
NO_CURSO 232-431 A200
CO_GRAU_ACADEMICO 432-439 F8
CO_MODALIDADE_ENSINO 440-447 F8
CO_NIVEL_ACADEMICO 448-455 F8
CO_ALUNO_CURSO 456-463 F8
CO_ALUNO 464-477 A14
CO_COR_RACA_ALUNO 478-485 F8
DS_COR_RACA_ALUNO 486-509 A24
IN_SEXO_ALUNO 510-517 F8
DS_SEXO_ALUNO 518-526 A9
NU_ANO_ALUNO_NASC 527-530 A4
NU_DIA_ALUNO_NASC 531-532 A2

```

NU_MES_ALUNO_NASC	533-534	A2
NU_IDADE_ALUNO	535-542	F8
CO_NACIONALIDADE_ALUNO	543-550	F8
CO_PAIS_ORIGEM_ALUNO	551-558	F8
CO_UF_NASCIMENTO	559-566	F8
CO_MUNICIPIO_NASCIMENTO	567-574	F8
CO_ALUNO_SITUACAO	575-582	F8
IN_ALUNO_DEFICIENCIA	583-590	F8
IN_DEF_AUDITIVA	591-598	F8
IN_DEF_FISICA	599-606	F8
IN_DEF_INTELECTUAL	607-614	F8
IN_DEF_MULTIPLA	615-622	F8
IN_DEF_SURDEZ	623-630	F8
IN_DEF_SURDOCEGUEIRA	631-638	F8
IN_DEF_BAIXA_VISAO	639-646	F8
IN_DEF_CEGUEIRA	647-654	F8
DT_INGRESSO_CURSO	655-664	A10
IN_ATIVIDADE_COMPLEMENTAR	665-672	F8
IN_RESERVA_VAGAS	673-680	F8
IN_FINANC_ESTUDANTIL	681-688	F8
IN_APOIO_SOCIAL	689-696	F8
CO_CURSO_POLO	697-704	F8
CO_TURNO_ALUNO	705-712	F8
IN_ING_VESTIBULAR	713-720	F8
IN_ING_ENEM	721-728	F8
IN_ING_OUTRA_FORMA_SELECAO	729-736	F8
IN_ING_CONVENIO_PEC_G	737-744	F8
IN_ING_OUTRA_FORMA	745-752	F8
IN_RESERVA_ETNICO	753-760	F8
IN_RESERVA_DEFICIENCIA	761-768	F8
IN_RESERVA_ENSINO_PUBLICO	769-776	F8
IN_RES_RENDA_FAMILIAR	777-784	F8
IN_RESERVA_OUTROS	785-792	F8
IN_FIN_REEMB_FIES	793-800	F8
IN_FIN_REEMB_ESTADUAL	801-808	F8
IN_FIN_REEMB_MUNICIPAL	809-816	F8
IN_FIN_REEMB_PROG_IES	817-824	F8
IN_FIN_REEMB_ENT_EXTERNA	825-832	F8
IN_FIN_REEMB_OUTRA	833-840	F8
IN_FIN_NAOREEMB_PROUNI_INTEGR	841-848	F8
IN_FIN_NAOREEMB_PROUNI_PARCIAL	849-856	F8
IN_FIN_NAOREEMB_ESTADUAL	857-864	F8
IN_FIN_NAOREEMB_MUNICIPAL	865-872	F8
IN_FIN_NAOREEMB_PROG_IES	873-880	F8
IN_FIN_NAOREEMB_ENT_EXTERNA	881-888	F8
IN_FIN_NAOREEMB_OUTRA	889-896	F8
IN_APOIO_ALIMENTACAO	897-904	F8
IN_APOIO_BOLSA_PERMANENCIA	905-912	F8
IN_APOIO_BOLSA_TRABALHO	913-920	F8
IN_APOIO_MATERIAL_DIDATICO	921-928	F8
IN_APOIO_MORADIA	929-936	F8
IN_APOIO_TRANSPORTE	937-944	F8
IN_COMPL_ESTAGIO	945-952	F8
IN_COMPL_EXTENSAO	953-960	F8

IN_COMPL_MONITORIA	961-968	F8
IN_COMPL_PESQUISA	969-976	F8
IN_BOLSA_ESTAGIO	977-984	F8
IN_BOLSA_EXTENSAO	985-992	F8
IN_BOLSA_MONITORIA	993-1000	F8
IN_BOLSA_PESQUISA	1001-1008	F8
TP_PROCEDE_EDUC_PUBLICA	1009-1016	F8
IN_MATRICULA	1017-1024	F8
IN_CONCLUINTE	1025-1032	F8
IN_INGRESSO	1033-1040	F8
IN_ING_PROCESSO_SELETIVO	1041-1048	F8
IN_ING_OUTRAS_FORMAS	1049-1056	F8
ANO_INGRESSO	1057-1060	A4
.		
CACHE.		
EXECUTE.		

APÊNDICES

A.6 SCRIPTS DA LINGUAGEM R DA APLICAÇÃO DO BIRTH

```
## Diretorio de Trabalho ##
setwd("E:/micro_censo_edu_superior2010/DADOS")
getwd()

## Carga dos pacotes utilizados ##
library(irtoys)
citation()
citation("irtoys")

## Importacao das respostas aos itens ##
R <- read.csv("E:/micro_censo_edu_superior2010/DADOS/R.csv",
  sep=",")
R1<-R[,3:18]
head(R1)
tail(R1)

## Analise CLASSICA - Dimensao Corpo Docente e Tutorial ##
R1.desc<-descript(R1)
R1.desc

## Analise TRI - Dimensao Corpo Docente e Tutorial ##
# MODELO ML2 - estimativas dos parametros
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# sem I108 I111
R1<-
R[,c("I101", "I102", "I103", "I104", "I105", "I106", "I107", "I109", "I
110", "I112", "I113", "I114", "I115", "I116")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# sem I108 I111 I116
R1<-
R[,c("I101", "I102", "I103", "I104", "I105", "I106", "I107", "I109", "I
110", "I112", "I113", "I114", "I115")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# sem I101 I108 I111 I116
R1<-
R[,c("I102", "I103", "I104", "I105", "I106", "I107", "I109", "I110", "I
112", "I113", "I114", "I115")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# sem I101 I108 I111 I115 I116
R1<-
R[,c("I102", "I103", "I104", "I105", "I106", "I107", "I109", "I110", "I
112", "I113", "I114")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
```

```

# sem I101 I108 I111 I113 I115 I116
R1<-
R[,c("I102","I103","I104","I105","I106","I107","I109","I110","I
112","I114")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# sem I101 I108 I111 I112 I113 I115 I116
R1<-
R[,c("I102","I103","I104","I105","I106","I107","I109","I110","I
114")]
R1.par<-est(R1, model="2PL", engine="bilog", nqp= 20)
R1.par
# MODELO ML2 - estimativas proficiencias - indice de
competitividade
R1.sco<-eap(R1,R1.par$est,qu=normal.qu())
head(R1.sco)
tail(R1.sco)

# Graficos
sco <- hist(R1.sco[,1], nclass = 4, plot = FALSE)
plot(sco, border = "dark blue", col = "light blue",
     main = "", xlab = "Índice de competitividade",
     ylab = "no. cursos", labels = TRUE,
     xlim = range(-3:3), ylim = range(0:800), )

sco
plot(irf(R1.par$est[1,]), main = "")
plot(irf(R1.par$est[2,]), main = "")
plot(irf(R1.par$est[3,]), main = "")
plot(irf(R1.par$est[4,]), main = "")
plot(irf(R1.par$est[5,]), main = "")
plot(irf(R1.par$est[6,]), main = "")
plot(irf(R1.par$est[7,]), main = "")
plot(irf(R1.par$est[8,]), main = "")
plot(irf(R1.par$est[9,]), main = "")
plot(iif(R1.par$est),label=TRUE, main = "")
plot(tif(R1.par$est),label=TRUE, main = "")

## Analise CLASSICA - modelo final ##
R1.desc<-descript(R1)
R1.desc
plot(R1.desc,type="b",includeFirstLast=TRUE)

```

A.7 LOG DA ANÁLISE CLÁSSICA DOS ITENS DA APLICAÇÃO DO BIRTH

BILOG-MG V3.0
 REV 19990104.1300
 BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL
 DISTRIBUTED BY
 SCIENTIFIC SOFTWARE INTERNATIONAL, INC.
 7383 N. LINCOLN AVENUE, SUITE 100
 CHICAGO, IL 60646
 (800) 247-6113
 (847) 675-0720
 WWW: <http://www.ssicentral.com>
 PROGRAM COPYRIGHT HELD BY SCIENTIFIC SOFTWARE INTERNATIONAL,
 INC. 2002
 DISTRIBUTION OR USE UNAUTHORIZED BY SSI, INC. IS PROHIBITED

*** BILOG-MG ITEM MAINTENANCE PROGRAM ***
 *** PHASE 1 ***

```
Running Bilog from R
>GLOBAL DFName = 'mymodel.dat',
  NPArm = 2,
  LOGistic
  SAVE;
```

FILE ASSIGNMENT AND DISPOSITION

```
=====
SUBJECT DATA INPUT FILE      MYMODEL.DAT
BILOG-MG MASTER DATA FILE    MF.DAT
                                WILL BE CREATED FROM DATA FILE
CALIBRATION DATA FILE        CF.DAT
                                WILL BE CREATED FROM DATA FILE
ITEM PARAMETERS FILE          IF.DAT
                                ZWILL BE CREATED THIS RUN
CASE SCALE-SCORE FILE         SF.DAT
CASE WEIGHTING                 NONE EMPLOYED
ITEM RESPONSE MODEL            2 PARAMETER LOGISTIC
                                LOGIT METRIC (I.E., D = 1.0)
```

```
>SAVE PARM = 'MYMODEL.BLMP';
BILOG-MG SAVE FILES
[OUTPUT FILES]
ITEM PARAMETERS FILE          MYMODEL.BLMP
>LENGTH NITems = (9);
```

TEST LENGTH SPECIFICATIONS

```
=====
MAIN TEST LENGTHS:           9
>INPUT NTotal = 9,
  NFName = 'myNF.ile'
  SAMple = 1581,
  NIDchar = 6;
```

DATA INPUT SPECIFICATIONS

```
=====
NUMBER OF FORMAT LINES                1
NUMBER OF ITEMS IN INPUT STREAM        9
NUMBER OF RESPONSE ALTERNATIVES       1000
NUMBER OF SUBJECT ID CHARACTERS        6
NUMBER OF GROUPS                       1
NUMBER OF TEST FORMS                   1
TYPE OF DATA                          SINGLE-SUBJECT DATA, NO CASE WEIGHTS
MAXIMUM SAMPLE SIZE FOR ITEM CALIBRATION 1581
ALL SUBJECTS INCLUDED IN RUN
>ITEMS INames = (ITEM0001(1)ITEM0009);
```

TEST SPECIFICATIONS

=====

```
>TEST1 TName = 'TEST',
  INumber = (1(1)9);
TEST NUMBER: 1    TEST NAME: TEST
NUMBER OF ITEMS: 9
```

ITEM NUMBER	ITEM NAME	ITEM NUMBER	ITEM NAME	ITEM NUMBER	ITEM NAME
1	ITEM1	4	ITEM4	7	ITEM7
2	ITEM2	5	ITEM5	8	ITEM8
3	ITEM3	6	ITEM6	9	ITEM9

FORM SPECIFICATIONS

=====

```
ITEMS READ ACCORDING TO SPECIFICATIONS ON THE ITEMS COMMAND
FORMAT FOR DATA INPUT IS:
(6A1,9A1)
```

KEY SPECIFICATIONS

=====

NOT PRESENTED KEY:

.....

OBSERVATION # 1 WEIGHT: 1.0000 ID : 000001

SUBTEST #: 1 TEST

GROUP #: 1

TRIED RIGHT

9.000 7.000

ITEM	1	2	3	4	5	6	7	8	9
TRIED	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RIGHT	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0

OBSERVATION # 2 WEIGHT: 1.0000 ID : 000002

SUBTEST #: 1 TEST

GROUP #: 1

TRIED RIGHT

9.000 6.000

ITEM	1	2	3	4	5	6	7	8	9
TRIED	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RIGHT	1.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	0.0

1581 OBSERVATIONS READ FROM FILE: MYMODEL.DAT
 1581 OBSERVATIONS WRITTEN TO FILE: MF.DAT

ITEM STATISTICS FOR SUBTEST TEST

ITEM	NAME	#TRIED	#RIGHT	PCT	ITEM*TEST CORRELATION		
					LOGIT	PEARSON	BISERIAL
1	ITEM1	1581.0	1149.0	72.7	-0.98	0.379	0.507
2	ITEM2	1581.0	577.0	36.5	0.55	0.466	0.596
3	ITEM3	1581.0	680.0	43.0	0.28	0.464	0.585
4	ITEM4	1581.0	305.0	19.3	1.43	0.396	0.570
5	ITEM5	1581.0	543.0	34.3	0.65	0.459	0.593
6	ITEM6	1581.0	452.0	28.6	0.92	0.342	0.455
7	ITEM7	1581.0	158.0	10.0	2.20	0.450	0.770
8	ITEM8	1581.0	629.0	39.8	0.41	0.634	0.805
9	ITEM9	1581.0	65.0	4.1	3.15	0.129	0.290

164 BYTES OF NUMERICAL WORKSPACE USED OF 2048000 AVAILABLE IN
 PHASE-1
 5124 BYTES OF CHARACTER WORKSPACE USED OF 409600 AVAILABLE IN
 PHASE-1

A.8 LOG DAS ESTIMATIVAS DOS PARÂMETROS DOS ITENS DA APLICAÇÃO DO BIRTH

BILOG-MG V3.0
 REV 19990329.1300
 BILOG-MG ITEM MAINTENANCE PROGRAM: LOGISTIC ITEM RESPONSE MODEL

*** BILOG-MG ITEM MAINTENANCE PROGRAM ***
 *** PHASE 2 ***

Running Bilog from R

>CALIB NQPt = 20,
 CYCles = 3000,
 NEWton = 0;

CALIBRATION PARAMETERS

```
=====
MAXIMUM NUMBER OF EM CYCLES:          3000
MAXIMUM NUMBER OF NEWTON CYCLES:      0
CONVERGENCE CRITERION:                0.0100
ACCELERATION CONSTANT:                1.0000
LATENT DISTRIBUTION:                  NORMAL PRIOR FOR EACH GROUP
PLOT EMPIRICAL VS. FITTED ICC'S:      NO
DATA HANDLING:                        DATA ON SCRATCH FILE
CONSTRAINT DISTRIBUTION ON SLOPES:     YES
CONSTRAINT DISTRIBUTION ON THRESHOLDS: NO
```

SOURCE OF ITEM CONSTRAINT DISTRIBUTION
 MEANS AND STANDARD DEVIATIONS: PROGRAM DEFAULTS

```
*****
CALIBRATION OF MAINTEST
TEST
*****
```

METHOD OF SOLUTION:
 EM CYCLES (MAXIMUM OF 3000)

QUADRATURE POINTS AND PRIOR WEIGHTS:

	1	2	3	4	5	
POINT	-0.4000E+01	-0.3579E+01	-0.3158E+01	-0.2737E+01	-	
0.2316E+01						
WEIGHT	0.5635E-04	0.2779E-03	0.1148E-02	0.3970E-02		
0.1150E-01						
	6	7	8	9	10	
POINT	-0.1895E+01	-0.1474E+01	-0.1053E+01	-0.6316E+00	-	
0.2105E+00						
WEIGHT	0.2791E-01	0.5671E-01	0.9653E-01	0.1376E+00		
0.1643E+00						
	11	12	13	14	15	

POINT	0.2105E+00	0.6316E+00	0.1053E+01	0.1474E+01
0.1895E+01				
WEIGHT	0.1643E+00	0.1376E+00	0.9653E-01	0.5671E-01
0.2791E-01				
	16	17	18	19
POINT	0.2316E+01	0.2737E+01	0.3158E+01	0.3579E+01
0.4000E+01				
WEIGHT	0.1150E-01	0.3970E-02	0.1148E-02	0.2779E-03
0.5635E-04				

CONSTRAINT DISTRIBUTIONS ON ITEM PARAMETERS
(THRESHOLDS, NORMAL; SLOPES, LOG-NORMAL; GUESSING, BETA)

ITEM	THRESHOLDS		SLOPES		ASYMPTOTES	
	MU	SIGMA	MU	SIGMA	ALPHA	BETA
ITEM1	-	-	1.000	1.649		
ITEM2	-	-	1.000	1.649		
ITEM3	-	-	1.000	1.649		
ITEM4	-	-	1.000	1.649		
ITEM5	-	-	1.000	1.649		
ITEM6	-	-	1.000	1.649		
ITEM7	-	-	1.000	1.649		
ITEM8	-	-	1.000	1.649		
ITEM9	-	-	1.000	1.649		

[E-M CYCLES]

-2 LOG LIKELIHOOD =	13370.296
CYCLE 1; LARGEST CHANGE=	0.47160
-2 LOG LIKELIHOOD =	13288.260
CYCLE 2; LARGEST CHANGE=	0.36511
-2 LOG LIKELIHOOD =	13266.013
CYCLE 3; LARGEST CHANGE=	0.29165
-2 LOG LIKELIHOOD =	13256.403
CYCLE 4; LARGEST CHANGE=	0.59209
-2 LOG LIKELIHOOD =	13246.692
CYCLE 5; LARGEST CHANGE=	0.34758
-2 LOG LIKELIHOOD =	13244.102
CYCLE 6; LARGEST CHANGE=	0.21637
-2 LOG LIKELIHOOD =	13242.954
CYCLE 7; LARGEST CHANGE=	0.21182
-2 LOG LIKELIHOOD =	13242.086
CYCLE 8; LARGEST CHANGE=	0.11349
-2 LOG LIKELIHOOD =	13241.700
CYCLE 9; LARGEST CHANGE=	0.06437
-2 LOG LIKELIHOOD =	13241.542
CYCLE 10; LARGEST CHANGE=	0.03961
-2 LOG LIKELIHOOD =	13241.426
CYCLE 11; LARGEST CHANGE=	0.02049
-2 LOG LIKELIHOOD =	13241.378
CYCLE 12; LARGEST CHANGE=	0.01308

-2 LOG LIKELIHOOD = 13241.350
 CYCLE 13; LARGEST CHANGE= 0.00863

SUBTEST TEST ; ITEM PARAMETERS AFTER CYCLE 13

ITEM	INTERCEPT	SLOPE	THRESHOLD	LOADING	ASYMPTOTE
ITEM1	1.326 0.074*	1.405 0.085*	-0.944 0.057*	0.815 0.049*	0.000 0.000*
ITEM2	-0.785 0.065*	1.540 0.087*	0.510 0.043*	0.839 0.047*	0.000 0.000*
ITEM3	-0.381 0.060*	1.385 0.080*	0.275 0.044*	0.811 0.047*	0.000 0.000*
ITEM4	-1.850 0.086*	1.288 0.088*	1.436 0.082*	0.790 0.054*	0.000 0.000*
ITEM5	-0.988 0.071*	1.788 0.098*	0.553 0.039*	0.873 0.048*	0.000 0.000*
ITEM6	-1.173 0.068*	1.224 0.078*	0.958 0.064*	0.774 0.050*	0.000 0.000*
ITEM7	-3.920 0.210*	2.495 0.173*	1.571 0.057*	0.928 0.064*	0.000 0.000*
ITEM8	-1.362 0.114*	5.026 0.285*	0.271 0.018*	0.981 0.056*	0.000 0.000*
ITEM9	-3.441 0.155*	0.798 0.122*	4.310 0.572*	0.624 0.095*	0.000 0.000*

* STANDARD ERROR

LARGEST CHANGE = 0.008635

NOTE: ITEM FIT CHI-SQUARES AND THEIR SUMS MAY BE UNRELIABLE
 FOR TESTS WITH LESS THAN 20 ITEMS

PARAMETER	MEAN	STN DEV
SLOPE	1.883	1.266
LOG(SLOPE)	0.493	0.519
THRESHOLD	0.993	1.447

QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D.:

	1	2	3	4	5
POINT	-0.4008E+01	-0.3586E+01	-0.3165E+01	-0.2743E+01	-
POSTERIOR	0.6310E-04	0.3137E-03	0.1282E-02	0.4366E-02	
	0.2321E+01	0.1232E-01			

6

7

8

9

10

POINT	-0.1899E+01	-0.1477E+01	-0.1055E+01	-0.6331E+00	-
0.2112E+00					
POSTERIOR	0.2877E-01	0.5593E-01	0.9231E-01	0.1328E+00	
0.1663E+00					

	11	12	13	14	15
POINT	0.2107E+00	0.6327E+00	0.1055E+01	0.1476E+01	
0.1898E+01					
POSTERIOR	0.1751E+00	0.1447E+00	0.8865E-01	0.4803E-01	
0.2734E-01					

	16	17	18	19	20
POINT	0.2320E+01	0.2742E+01	0.3164E+01	0.3586E+01	
0.4008E+01					
POSTERIOR	0.1400E-01	0.5578E-02	0.1716E-02	0.4173E-03	
0.8186E-04					

MEAN	0.00000
S.D.	1.00000

15428 BYTES OF NUMERICAL WORKSPACE USED OF 2048000 AVAILABLE IN PHASE-2

3020 BYTES OF CHARACTER WORKSPACE USED OF 409600 AVAILABLE IN PHASE-2

11/13/2012 19:20:43