

*Teoria de Resposta ao Item (TRI) na análise de pesquisa de desempenho na gestão da qualidade em uma siderúrgica.*

**Tatiana Rodrigues Quintana – [trquintana@gmail.com](mailto:trquintana@gmail.com) - UFRGS**

**Márcia Echeveste – [echeveste@producao.ufrgs.br](mailto:echeveste@producao.ufrgs.br) – UFRGS**

Resumo: Um indicador muito utilizado pelas empresas na gestão da qualidade são os índices de satisfação obtidos através de pesquisas ao consumidor, contudo este indicador deve estar próximo das verdadeiras medidas para retratar a realidade da empresa. Neste contexto, o fator motivador deste estudo partiu da análise dos dados de SAC do setor comercial de uma siderúrgica que utiliza uma pesquisa de satisfação para avaliar seu serviço de vendas. O setor apontava um expressivo número reclamações para o serviço de atendimento ao consumidor, contudo a pesquisa apresentava um alto percentual de satisfeitos. Este artigo propõe uma nova forma de análise para a pesquisa utilizando a teoria de resposta ao item (TRI) como alternativa à teoria clássica. Os resultados demonstraram que a TRI possui benefícios perante a teoria clássica, gerando análises mais detalhadas, com indicadores mais coerentes entre as reclamações e o percentual de satisfação, contribuindo à empresa na sua gestão da qualidade.

Palavras chave: teoria de resposta ao item, gestão da qualidade, pesquisa de satisfação.

*Abstract: An indicator commonly used by companies in the management of quality are the satisfaction obtained through consumer surveys, however this indicator should be close to the real action to portray the reality of the company. In this context, the motivating factor of this study was the analysis of data from the SAC the commercial sector of a steel plant that uses a satisfaction survey to assess their sales service. The industry indicated a significant number of complaints to the service customer service, yet the survey had a high percentage of satisfied. This article proposes a new form of analysis for research using item response theory (IRT) as an alternative to the classical theory. The results showed that the IRT has benefits before the classical theory, generating more detailed analysis, with most indicators consistent with the claims and the percentage of satisfaction, contributing to the company in its quality management.*

*Keywords: item response theory, quality management, satisfaction survey.*

## 1. Introdução

Na gestão da qualidade em produtos e serviços, pesquisas de qualidade de avaliação de serviço e vendas são importantes fontes de informações para o gerenciamento e proposição de soluções integradas para produtos e serviços. Destaca-se que ações podem ser disparadas a partir destes resultados, como a priorização de projetos de melhoria relacionados à análise do ciclo de vida de produtos, tais ações incluem: melhorias nos produtos já existentes, idéias e sugestões de novos produtos, assim como decisões de retirada de linha. Estas informações impactam diretamente na gestão do portfólio de produtos bem como em ações localizadas de melhoria da qualidade como treinamento de operadores e análise de solução de problemas.

Nas grandes empresas utilizam-se pesquisas de satisfação de cliente como base para impulsionar projetos de melhoria. Contudo, as pesquisas muitas vezes são elaboradas de forma a não trazer resultados que sejam úteis para o controle da qualidade de produtos e serviços. Segundo Juran (1993), a aplicabilidade da qualidade total está ligada a um bom indicador, que realmente represente o que esta se buscando. De nada adianta traçar um plano para melhoria da qualidade se o indicador usado é falho, se ele não representa a real deficiência ou necessidade da empresa, desta forma esta busca pela melhoria da qualidade através dele levará a erros, perdas financeiras e de tempo (MULLER, 2003).

Em grande parte das áreas da engenharia de produção são utilizadas escalas de medida, tanto para medir clima da organização, como controlar indicadores de desempenho, nas questões ambientais, qualidade, maturidade das organizações, no processo de inovação, na identificação de fatores críticos na gestão de intangíveis, satisfação de clientes, etc. Quando bem elaboradas, estas escalas nos permitem identificar erros no processo, o que pode apontar oportunidades de melhoria e poder de comparabilidade (benchmarking), por exemplo, (TEZZA, 2009). Neste contexto, pesquisas de marketing são amplamente utilizadas como indicadores de necessidades de clientes como também indicadores de deficiências de uma empresa. A elaboração de instrumentos de pesquisa fidedignos pode trazer elementos confiáveis para a tomada de decisão, e na maioria das empresas, esses instrumentos são baseados em questionários elaborados segundo a Teoria Clássica de Medida (TCM).

A TCM é uma forma de análise que utiliza as respostas dadas a um questionário, após a sua aplicação, considerando todos os itens do questionário idênticos, não levando em conta seus parâmetros de dificuldade e/ou discriminação na avaliação dos respondentes. Além disso, ela não tem como foco a análise de cada item, e sim o questionário como um todo. Contudo, recentemente, em alguns instrumentos de pesquisa como, por exemplo, pesquisa de satisfação, propõem como alternativa o uso da Teoria de Resposta ao Item (TRI), por que esta avalia de forma mais aprimorada o nível de satisfação de cada indivíduo participante. Além disto, ela leva em consideração os parâmetros dos itens, o que não ocorre com a TCM. A TRI também faz uma validação do instrumento de pesquisa de uma forma diferente da TCM, o que pode servir para mostrar que itens não estão sendo bem usados no questionário, possibilitando o aprimoramento deste.

Assim sendo, a proposta de estudo é aplicar a Teoria de Resposta ao Item à engenharia, usando-se para isso a análise de dados da qualidade percebida de clientes quanto à qualidade do serviço de vendas como alternativa à Teoria Clássica de Medida na pesquisa, comparando os resultados dessas teorias. Para tanto, será estudado o caso real de uma empresa multinacional do ramo siderúrgico que utiliza um questionário de avaliação do serviço de venda prestado pelo setor comercial, no qual tem como base de gestão a Gestão da Qualidade Total. A justificativa para o estudo é que como a análise feita pela TRI é item a item, pode-se “visualizar” quais itens estão funcionando bem no questionário e quais devem ser modificados. A TRI também pode mostrar para quais níveis de satisfação do cliente que o questionário funciona melhor, uma vez que a TRI verifica que para cada nível de satisfação do cliente o teste vai funcionar melhor ou pior, já a TCM pré supõe que a pesquisa funciona tão bem para avaliar um cliente insatisfeito quanto para avaliar um satisfeito. Contudo, neste estudo, a TRI será usada como uma alternativa para validar e refinar o instrumento de pesquisa já utilizado pela empresa objeto de estudo, fazendo com que a avaliação da qualidade seja aprimorada para uma melhor gestão da empresa.

## 2. Referencial teórico

“Nas pesquisas empíricas, em particular nas organizacionais, a confiabilidade das conclusões está diretamente relacionada à validação do instrumento de medição” (WELLIANDRE, 2002, p.13). Ou seja, identificar problemas, entendê-los e propor melhorias segundo os princípios da Gestão da Qualidade Total (GQT) parte de um levantamento de dados eficaz. Por tanto, como se refere Drucker (1995), o método escolhido para esse levantamento precisa estar adequado ao contexto e a confiabilidade das informações provindas deste. Uma empresa que preza pela Gestão da Qualidade Total (GQT) utilizando ferramentas como pesquisa de marketing para medir a satisfação de seus clientes deve estar atenta aos seguintes questionamentos:

- i) Como modelar uma pesquisa que responda com confiabilidade as dúvidas da empresa sobre a aplicação de seus serviços?
- ii) Como mensurar qualitativamente uma pesquisa de satisfação em serviços que por sua natureza são intangíveis (VARGAS ET AL., 2007)?
- iii) Como confiar que a resposta do cliente realmente demonstra seu grau de satisfação e não esteja contaminada por outros fatores?

Testes e questionários são instrumentos de medida úteis no auxílio nas tomadas de decisão dos profissionais de diversas áreas. Quando se produz um questionário objetiva-se através deste documento se obter informações dos comportamentos que identifiquem o aspecto latente aos respondentes (que não seja identificado através de observações). Este comportamento é quantificado de forma que se obtenha um escore numérico (LORD, 1980). Da produção do questionário podem ser utilizados vários procedimentos e formas nesse processo e no caso de medidas objetivas (questionários), pode-se fundamentar esta avaliação em dois modelos: a Teoria Clássica de medida (TCM) e a Teoria de Resposta ao Item (TRI).

Na Teoria Clássica de Medida o resultado da mensuração através do escore do questionário necessita especificamente de uma comparação. Um exemplo é medir a altura de uma pessoa, para isso o comprimento da pessoa é comparado a uma escala, nesse caso a escala métrica. Para haver uma comparação devem ser especificados o padrão com o qual o escore é comparado e uma base numérica de comparação (ordem, comparação, diferença etc.). No caso da altura, como no exemplo, o padrão de comparação é uma escala métrica e a base numérica é de relação (EMBRETSON E

REISE, 2000). Na Teoria Clássica de Medida o padrão de comparações entre o escore é a base de dados.

Já na TRI leva-se em conta os níveis de traços latentes, ou seja, supõem-se que exista no sujeito respondente um traço (uma característica individual que determina a forma como responder aos itens de um teste) que possui individualmente com os itens utilizados uma relação probabilística (FLETCHER, 1994). Desta forma se considera que os parâmetros de cada item são estimados de forma independente dos demais itens do teste ou questionário. Contudo, o escore do questionário é obtido em função das respostas do respondente a cada item. Também dessa forma é possível constatar se os respondentes são mais ou menos hábeis ou se os itens do questionário devem ser considerados mais fáceis ou mais difíceis, já que na TRI coloca-se pessoas e itens em uma escala comum.

## 2.1 Método TCM

Esta teoria foi lançada em 1808 e inicialmente foi utilizada para avaliações psicológicas em pesquisas com pacientes em 1859. Atualmente a TCM é utilizada com método para calcular o escore da grande maioria das universidades do país e amplamente difundida para mensurar pesquisas de opinião.

Esta teoria é focada na explicação do resultado final de um teste ou questionário somando os escores brutos obtidos em seus itens. Em outras palavras, da pesquisa produzida se obtém das respostas dos clientes um escore matemático dito verdadeiro e um erro considerado aleatório. A equação (1) que representa esta teoria foi descrita em 1992 por Hayes:

$$X = T + e \quad (1)$$

Onde:

X é o escore observado

T o escore verdadeiro

e o erro aleatório.

Escore verdadeiro T é a soma das respostas dadas aos itens da pesquisa. O erro e, que pode acontecer na pesquisa são erros como, por exemplo: o respondente errar a resposta; o avaliador passar os dados errados e o avaliador arredondar os escores erroneamente. Este erro é dito como aleatório exatamente pela imprecisão dos erros que

podem acontecer na aplicação do questionário. Ainda também o que acontece na TCM é que o erro de medida  $E$  quando pequeno pode significar que os escores de  $X$  são confiáveis em relação ao escore verdadeiro  $T$ . Porém este erro aleatório não mensurado nesta teoria é abrangente e, segundo Vianna (1987) pode significar ainda que haja no questionário produzido a impossibilidade de um número infinito de questões, não possibilidade da aplicabilidade do questionário infinitas vezes ou o questionário não é aplicável a qualquer cliente ou organização, o que o torna muitas vezes ineficaz ou não confiável. Outras limitações da TCM são que, para esta teoria, os itens de um questionário são de igual valor entre si e os respondentes vistos como seres idênticos.

## 2.2 Método TRI

A Teoria de Resposta ao Item (TRI) surgiu no início do século 20 nos estudos de Thurstone que desenvolveu um método de medida estatístico denominado Lei dos Julgamentos Comparativos, posteriormente este estudo foi aprimorado e difundido por Frederic Lord nos anos de 1952 e 1953, onde ganhou o nome de Teoria de Resposta ao Item. A TRI traz uma nova proposta estatística sem entrar em contradição com os princípios da psicometria: a de analisar de forma centrada os itens, superando desta forma as limitações da teoria clássica (NUNES et. al., 2008).

A psicologia rege que o comportamento dos humanos vem da consequência de diversos processos hipotéticos intitulados de traços latentes (LAZERSFELD, 1950). Desta forma, a TRI demonstra a relação entre os comportamentos (denominadas variáveis observáveis) e os traços latentes (variáveis hipotéticas) utilizando para isso uma equação matemática chamada de equação logística. Por isso, nessa modelagem matemática são representadas a probabilidade de determinada resposta a um item ser apontada em função dos parâmetros que caracterizam este item e do nível do respondente quanto ao traço latente que está sendo medido, o que a torna mais completa que a TCM que “desprezaria” o traço latentes, ou seja, a diversificação do comportamento humano. E conforme Pasquali e Primi (2003), além do uso do traço latente, a superioridade da TRI é comprovada pelo fato desta atender três limitações da TCM, que segundo Hambleton, Swaminathan e Rogers (1991 apud PAQUALI E PRIMI, 2003), são:

- (i) Na TCM os parâmetros de discriminação e dificuldade dos itens dependem da quantidade da amostra utilizada na pesquisa. Entretanto, por muitas vezes existe a dificuldade de se conseguir um número representativo para a amostra, o que pode tornar o estudo não confiável. Na TRI, este cálculo dos parâmetros de discriminação e dificuldade de um item não dependem do tamanho da amostra;
- (ii) A precisão na TCM é calculada pela correlação dos escores finais de formas paralelas de um teste e o erro de medida é tomado como idêntico entre os sujeitos da amostra, porém as pessoas são diferentes, com pensamentos e habilidades distintas. Dentro da TRI a precisão da pesquisa não exige que as formas sejam rigorosamente paralelas. E não existe a necessidade que o erro de medida seja suposto igual para todos os sujeitos da amostra.
- (iii) Todas as conclusões da TCM são baseadas no teste em geral. Os itens são analisados de forma abrangente. Entretanto, a TRI analisa item a item, de forma particular independente dos demais itens da pesquisa levando em conta sua função individual no instrumento de medida.

Além da área psicométrica e educacional, no Brasil a TRI vem sendo utilizada em outras áreas: serviços (COSTA, 2001), qualidade de vida (GUEWEHR, 2007), avaliação de intangíveis (VARGAS ET AL, 2007), gestão pela qualidade total (ALEXANDRE et al., 2002) e usabilidade em sites de e-commerce (TEZZA, 2009). Para adequar-se a TRI em pesquisas de marketing este estudo testa o modelo politômico. Na qual é caracterizado por um conjunto ordenado de respostas e não pela resposta binária ( 0 e 1), errado ou certo o que caracterizaria o modelo dicotômico. Este modelo de respostas, chamado de graduado, foi desenvolvido por Samejima (1969). Para um estudo mais aprofundado das modelagens utilizadas pela TRI recomenda-se a leitura de VAN DER LINDEN, 1997.

### **2.2.1 Modelo de respostas graduadas (MRG).**

Assume-se no modelo de Samejima (1969) que categorias de respostas de um item podem ser ordenadas entre si. Supõe-se que os escores das categorias de um determinado item  $i$  estão distribuídos em uma ordem crescente considerada por  $k = 0, 1, \dots, m_i$  onde  $(m_i + 1)$  representa o número de categorias do  $i$ -ésimo item. Então a

análise é realizada da seguinte forma: a probabilidade de um respondente  $j$  escolher uma específica categoria ou alguma mais alta do item  $i$  é representada pela equação (2).

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-D a_i(\theta_j - b_{i,k})}} \quad (2)$$

Onde:

$i$ : Número de itens do teste;

$j$ : Número de respondentes;

$\theta_j$ : Atitudes em relação à estatística do  $j$ -ésimo indivíduo (traço latente);

$a_i$ : Parâmetro discriminação do item  $i$ .

$b_{i,k}$ : Parâmetro de dificuldade da  $k$ -ésima categoria do item  $i$ . Pode ser particionado em dois termos, sendo  $b_i$  o parâmetro de dificuldade geral do item.

$D$ : Constante equivalente a 1 para a curva logística ou a 1,7 para obter a aproximação para a distribuição normal.

Em um modelo de, por exemplo, dois parâmetros, com  $i=1, \dots, I$ ;  $j=1, \dots, n$ ;  $b_{i,k}$ , é o parâmetro de dificuldade da  $k$ -ésima categoria do item  $i$ .

Ainda por definição do modelo politômico tem se que:

$$b_{i,1} \leq b_{i,2} \leq \dots \leq b_{i,m_i}$$

Que representa o ordenamento entre os níveis de dificuldade das categorias de um dado item.

Já a probabilidade de um indivíduo dito  $j$  obter um escore  $k$  no item  $i$  é dada por:

$$P_{i,k}(\theta_j) = P_{i,k}^*(\theta_j) - P_{i,k+1}^*(\theta_j) \quad (3)$$

Admitindo-se que:

$$P_{i,0}^*(\theta_j) = 1 \text{ e } P_{i,m_i+1}^*(\theta_j) = 0 \quad (4)$$

Tem-se:

$$P_{i,0} = P_{i,0}^*(\theta_j) - P_{i,1}^*(\theta_j) = 1 - P_{i,1}^*(\theta_j) \quad (5)$$

$$P_{i,1} = P_{i,1}^*(\theta_j) - P_{i,2}^*(\theta_j) \quad (6)$$

$$P_{i,m} = P_{i,m}^*(\theta_j) - P_{i,m+1}^*(\theta_j) = P_{i,m}^*(\theta_j) \quad (7)$$

Generalizando:

$$P_{i,k} = P_{i,k}^*(\theta_j) - P_{i,k+1}^*(\theta_j) \geq 0 \quad (8)$$

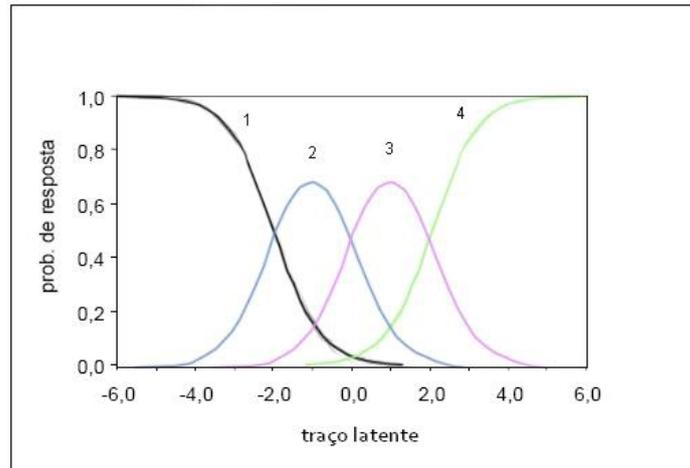
A partir dessas análises obtém-se a equação logística do modelo de respostas graduadas, conforme equação (9).

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k+1})}}, \quad (9)$$

### 2.2.2. Interpretação dos coeficientes e da Curva característica de resposta

Na TRI o parâmetro de discriminação representa a capacidade do item de diferenciar indivíduos com diferentes níveis do traço latente. Já o parâmetro de dificuldade indica que um item é considerado tanto mais difícil quanto maior for o nível de traço latente que o sujeito necessita apresentar para aceitá-lo.

No gráfico produzido no modelo politômico as curvas geradas são denominadas Curvas de Categoria de Resposta. A figura 1 representa graficamente as quatro curvas diferentes geradas por um item com quatro possibilidades de resposta, cada curva representa uma alternativa. No eixo horizontal são representados os valores do traço latente que é medido pelo item e no eixo vertical, a probabilidade de escolha das alternativas, que varia de 0 a 1.



**Figura 1- Exemplo politômico de curvas de categoria de resposta**  
**Fonte: TURIK (2010)**

Interpreta-se pela leitura do gráfico, que respondentes com nível de traço latente até -2 possuem uma maior probabilidade de optar pela primeira categoria de resposta. Respondentes com nível de traço latente entre -2 e 0 têm maior probabilidade de responderem à segunda categoria de resposta. Já Respondentes com nível de traço latente entre 0 e 2 a maior probabilidade é que escolham a categoria três. Por conseguinte, para respondentes com nível de traço latente acima de 2 a maior probabilidade é que escolham a última categoria de resposta (TURIK, 2010). Pode-se perceber uma associação entre o acréscimo na dimensão subjacente e um correspondente incremento na probabilidade de escolha de respostas com pontuações mais altas na escala, como se cada resposta indicasse um passo mais adiante na escala subjacente. Demonstrando dessa forma que à medida que se avança no nível do traço latente medido maior a probabilidade sucessiva que pontuações mais altas sejam escolhidas (NUNES et al., 2008).

### 2.2.3 Softwares para análise da TRI

Para as análises dos dados e representações gráficas são utilizados softwares como o Bimain, Bilog, Microcat, Parscale dependendo do tipo de item e do tipo de análise que se deseja (estimação de parâmetros TRI e da Teoria Clássica e cálculo de estatísticas para verificação das estimativas). Para a análise da unidimensionalidade dos utiliza-se Testfact e Microfact. Neste estudo de utilizará o software Parscale 4.1 por este

ser mais acessível, sendo disponibilizado pelos desenvolvedores na *internet* e de fácil programação.

### **3. Metodologia de pesquisa**

Para a comparação dos métodos se utilizará uma pesquisa de satisfação de clientes de grande siderúrgica multinacional onde são realizadas semanalmente reuniões com a equipe de vendas do setor comercial para constatar problemas, analisar indicadores de desempenho e satisfação, como propor mudanças ou melhorias no grupo. Estes indicadores interferem diretamente na escolha de estratégia de vendas da empresa. Nestas reuniões os resultados de pesquisas, incluindo a de satisfação de clientes, são apresentados e discutidos pelo grupo. Contudo, tornou-se freqüente nestas reuniões discussões sobre o crescente índice de reclamações referentes ao serviço de vendas recebido pelo serviço de SAC (serviço de atendimento ao consumidor) desta empresa. Porém, o aumento crescente deste número não estava sendo representado pelos índices apresentados na pesquisa de satisfação de clientes que se mantinha a um nível satisfatório não demonstrando problemas neste setor. Esta diferença de perspectivas pode ser um caso de indicador falho, questionando-se então se a pesquisa está condizendo com a realidade das necessidades dos clientes e deficiências da empresa.

A gestão da qualidade torna-se necessária visto que ela determina as diretrizes da empresa. Seu alinhamento com a pesquisa de satisfação deve ser preciso, pois este sendo um indicador de desempenho não pode ser falho na busca por melhorias no setor.

Após a conclusão que o indicador pode ser falho, parte-se para a busca de sua melhoria. Visto que em uma pesquisa de opinião há dois principais pontos que podem estar equivocados: as perguntas realizadas na pesquisa ou a forma de análise desta. Entretanto como esta pesquisa realizada em todas os setores comerciais desta multinacional, tanto em matrizes como filiais, de acordo com a visão, missão e valores da empresa, neste estudo se optou questionar a forma de sua análise, não o seu conteúdo.

Partindo-se para pesquisa de fórmulas de análises alternativas em periódicos e livros nacionais e internacionais constatou-se em segmentos de pesquisa, incluindo a satisfação de cliente, a grande procura pela TRI substituindo a TCM por sua confiabilidade mais precisa. Após pesquisa sobre a TRI e seus benefícios perante a

TCM, escolhe-se esse método como alternativa de análise e opta-se por comparar as duas em uma das pesquisas utilizadas pelo setor comercial da empresa.

Para tanto foi escolhida a pesquisa realizada em abril do ano de 2010, pois o mês de abril é considerado estável no seu volume de vendas, no decorrer dos anos não demonstrando grandes alterações mesmo com a sazonalidade do mercado siderúrgico. Nesta pesquisa foram entrevistados 1043 clientes que utilizaram o serviço de vendas da empresa neste mês. Trata-se de uma pesquisa quantitativa visto que as informações são colhidas por meio de um questionário padronizado e uniformizado, no qual os clientes marcam seu índice de satisfação para cada item declarado. A figura 2 demonstra o número de vendas comparados ao número de reclamações do setor comercial.

A escolha desta pesquisa como estudo foi motivada pela equipe de vendas do setor comercial da empresa em questão, na qual gerencia indicadores de melhorias baseados em pesquisas, tais como de satisfação dos clientes, de satisfação dos empregados, de cargos e salários, como exemplos. Estes indicadores impactam na premiação ou deleção de funcionários como também na estratégia de vendas do grupo no qual este empresa esta inserida.

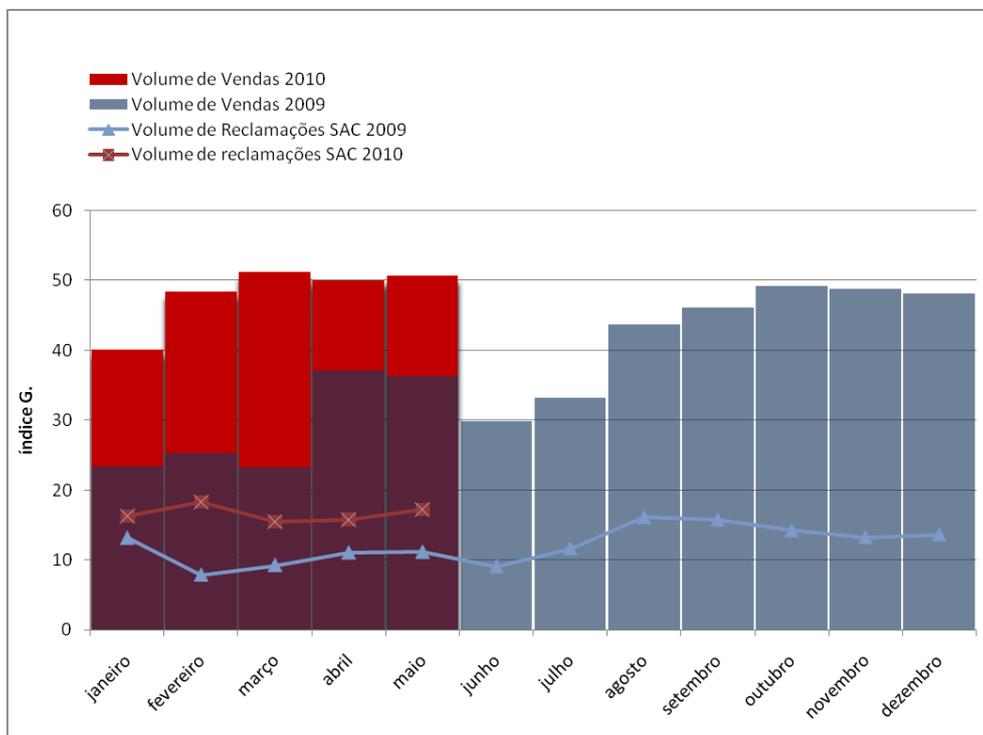


Figura 2 - Comparação do volume de vendas (índice G) com percentual de reclamações no SAC.

Finalizando, para compararem-se os dois métodos de análise, utiliza-se o banco de dados atual, analisando primeiramente pelo método da TCM, expondo passo a passo de cada etapa desta análise e demonstrando seus resultados. Após parte-se para a análise da pesquisa utilizando a TRI, descrevendo detalhadamente suas etapas até os resultados finais. Para concluir o estudo, os resultados das duas formas de análise serão comparados ao índice de reclamações do SAC demonstrado na Figura 2, discutindo-se, após, os resultados apresentados.

#### **4. Aplicação e Resultados:**

A pesquisa de satisfação de clientes da área comercial da empresa estudada compreendeu 1043 entrevistados, clientes pessoa física a jurídica, com idades variando entre 22 a 74 anos no mês de Abril de 2010. Este questionário é feito com todos os clientes, em material próprio da empresa, sendo entregue ao final do atendimento em uma urna, sem identificação formal, constando apenas a idade, tipo e quantidade de compra do cliente. Atendendo ao pedido da empresa o questionário não será anexado a esse trabalho, porém o banco de dados e suas questões disponibilizaram o estudo deste.

A pesquisa apresenta 6 questões aos clientes: (i) Quanto ao atendimento, a Comercial superou suas expectativas ?; (ii) Quanto as informações, os vendedores as passaram claramente ?; (iii) Quanto a agilidade do atendimento?; (iv) Quanto a equipe, demonstrou-se conhecimento técnico ?; (v) Quanto a cordialidade dos atendentes?; (vi) Quanto ao suporte pós venda foi satisfatório?

As questões devem ser respondidas através de escala similar a likert, onde a cada resposta se atribui um valor: de 1 (não concordo totalmente); 2 (concordo parcialmente); 3 (indiferente); 4 (concordo parcialmente) a 5 (concordo totalmente).

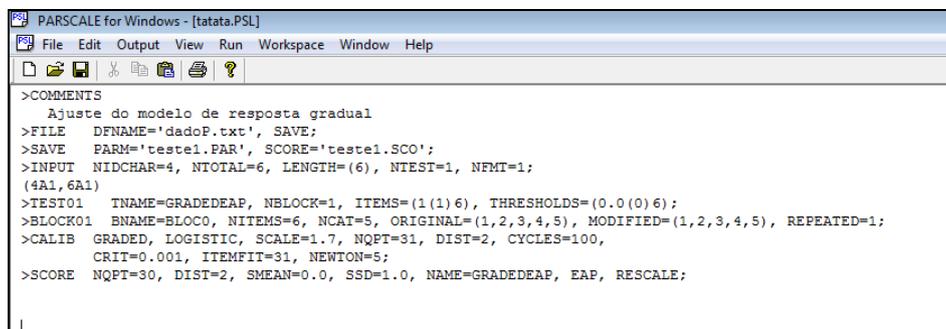
##### **4.1 Análise atual:**

Para a empresa um cliente é considerado satisfeito com o item questionado quando atribui nota superior ou igual a 3 a este. Analisando o banco de dados dos 1043 respondentes a 5 questões, a empresa identifica das 6258 pontuações atribuídas notas de 1 a 5. Dessas, 4979 respostas obtiveram valor igual ou acima de 3. Concluindo então

que 80% de seus clientes do mês de abril estão satisfeitos com o serviço de vendas prestado.

## 4.2 Análise TRI.

Para aplicação do Modelo de Resposta Gradual de Samejima (1969) da Teoria de Resposta ao Item a pesquisa foi utilizado o software PARSCALE, ilustrado na figura 3. Primeiramente é necessário programar o uso para uma análise de respostas gradual, para o método de Samejima politômico. Após insere-se o banco de dados da pesquisa estudada.



```
PARSCALE for Windows - [tatata.PSL]
File Edit Output View Run Workspace Window Help
>COMMENTS
  Ajuste do modelo de resposta gradual
>FILE  DFNAME='dadoP.txt', SAVE;
>SAVE  PARM='teste1.PAR', SCORE='teste1.SCO';
>INPUT NIDCHAR=4, NTOTAL=6, LENGTH=(6), NTEST=1, NFMT=1;
(4A1,6A1)
>TEST01  TNAME=GRADEDEAP, NBLOCK=1, ITEMS=(1(1)6), THRESHOLDS=(0.0(0)6);
>BLOCK01  BNAME=BLOC0, NITEMS=6, NCAT=5, ORIGINAL=(1,2,3,4,5), MODIFIED=(1,2,3,4,5), REPEATED=1;
>CALIB  GRADED, LOGISTIC, SCALE=1.7, NQPT=31, DIST=2, CYCLES=100,
        CRIT=0.001, ITEMFIT=31, NEWTON=5;
>SCORE  NQPT=30, DIST=2, SMEAN=0.0, SSD=1.0, NAME=GRADEDEAP, EAP, RESCALE;
```

Figura 3- Programando PARSCALE

Após rodado sua primeira fase o programa gera escores respondente a respondente e dados dos itens. O programa analisa primeiro o comportamento dos respondentes juntos aos itens, analisa item a item e a por fim análise respondente a respondente. Para a análise TRI, os resultados são analisados a partir da fórmula do modelo de Samejima descrito anteriormente na equação (9).

### 4.2.1 Análise de Parâmetros:

#### 4.2.1.1 Traço Latente:

Analisando primeiro o parâmetro  $\theta$ , que representa o traço latente. Este parâmetro pode assumir, teoricamente, qualquer valor entre  $-\infty$  e  $+\infty$ . Portanto é estabelecida uma origem e uma unidade de medida para definição da escala de traços latentes. Neste estudo utilizou-se uma escala com média igual à zero ( $\mu = 0$ ) para representar a média dos traços latentes dos respondentes do questionário e desvio-padrão igual a um ( $\sigma = 1$ ). As estimativas identificadas dos parâmetros do traço latente apresentaram uma variação de -3,10 a 3,44, distribuídos de acordo com a figura 4.

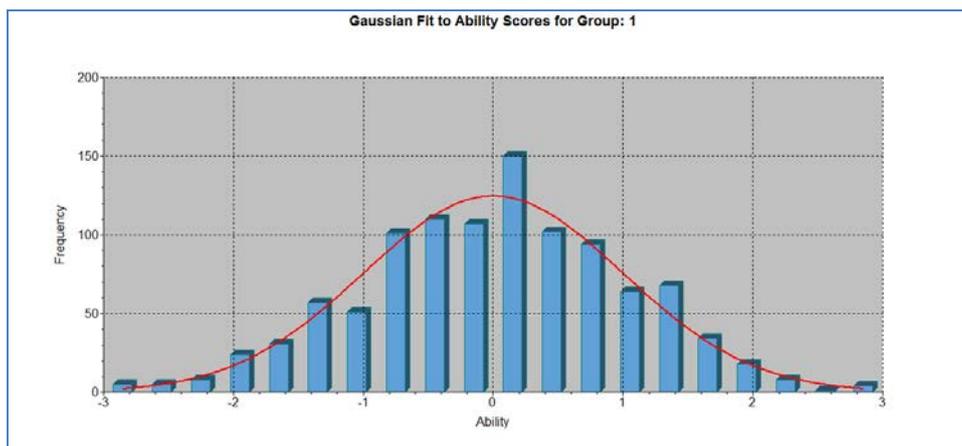


Figura 4: Histograma dos Escores Latentes

#### 4.2.1.2 Poder de Discriminação:

Uma importante característica que reflete a qualidade de um item é o poder de conseguir discriminar sujeitos com magnitudes diferentes no traço latente avaliado. A investigação de cada um dos itens da escala da pesquisa estudada revelou os parâmetros de discriminação ( $a_i$ ) variando entre 0.403 (item 4) e 0.692 (item 3). A literatura sugere categorias de discriminação, propondo como fracas até  $a_i = 1,34$ , moderadas entre 1,35 e 1,69 e altas a partir deste número (ALBUQUERQUE e TRÓCCOLI, 2004). Segundo a tabela gerada por PARSCALE os itens da pesquisa demonstram discriminação muito baixa, estando todos abaixo de  $a_i = 1,34$  como ilustra a tabela 1.

Item	$a_i$
1.Quanto ao atendimento, a Comercial superou suas expectativas	0.463
2.Quanto as informações, os vendedores as passaram claramente	0.479
3. Quanto a agilidade do atendimento	0.692
4. Quanto a equipe, demonstrou-se conhecimento técnico	0.403
5. Quanto a cordialidade dos atendentes	0.584
6. Quanto ao suporte pós venda foi satisfatório	0.519

Tabela 1: Discriminação dos itens

#### 4.2.1.3 Parâmetro de dificuldade:

As estimativas do parâmetro de dificuldade do item ( $b_i$ ) estão na mesma escala que as estimativas do traço latente ( $\theta$ ). Os valores negativos obtidos gerados na tabela 3 apontam que todas as questões são relativamente fáceis para serem respondidas, ou seja, que nenhum item requer que o respondente precise estar muito favorável para estar de acordo com este. O valor mais alto é do item 2 (Quanto as informações, os vendedores as passaram claramente). Este resultado aponta que o item requer atitudes mais intensamente favoráveis para que o respondente esteja de acordo com este item ( $b_i = -0.291$ ). Com base nisso, pode-se concluir que existe falha na geração de informações para os clientes, segundo a percepção do respondente, tendo que a empresa rever essa dinâmica. Por outro lado, o valor mais baixo ( $b_i = -7.911$ ) é do item 5 (Quanto a cordialidade dos atendentes:). Isso demonstra que este é respondido com maior clareza e convicção, os respondentes já estão pré dispostos a concordarem com o item. Isso pode demonstrar um item ruim para a pesquisa pois tendencia sua resposta. Porém demonstra uma questão a não ser revista pela empresa já que seus clientes naturalmente concordam com a cordialidade dos vendedores.

Item	$b_i$
1.Quanto ao atendimento, a Comercial superou suas expectativas	-0.646
2.Quanto as informações, os vendedores as passaram claramente:	-0.291
3. Quanto a agilidade do atendimento	-5.924
4. Quanto a equipe, demonstrou-se conhecimento técnico	-0.649
5. Quanto a cordialidade dos atendentes	-7.911
6. Quanto ao suporte pós venda foi satisfatório	-6.567

Tabela 2: Parâmetro de dificuldade dos itens

#### 4.2.2 Curva de informação do teste:

Uma importante informação gerada por PARSCALE é a Curva de Informação do Teste (Figura 5). Os itens envolvidos na pesquisa devem ser suficientes para que o instrumento possa ser informativo nos diferentes níveis da escala de escore latente. Considerando os 6 itens analisou-se a Curva de Informação do Teste (Figura 5). A curva em linha contínua mostra que a escala fornece maior quantidade de informações para respondentes com intensidade de traços latentes variando de 2,10 a 2,50. Contudo, a curva em linha tracejada representa que as atitudes são mais precisamente mensuradas, com menor erro, no intervalo, de 2,20 a 2,63. Comparando esta informação com o

gráfico das frequências de traços latentes (figura 4), nota-se que a população de respondentes neste intervalo de 2,10 a 2,63 é bastante reduzida frente ao todo. O que leva a conclusão que as informações geradas pelos itens não são suficientes para os diferentes níveis de traços latentes.

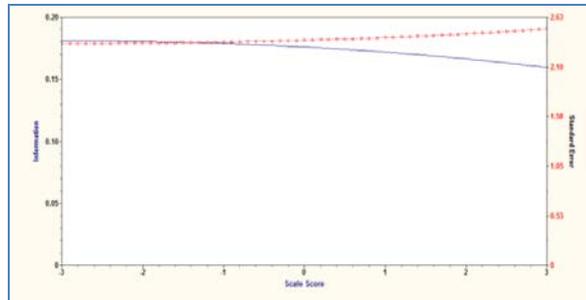
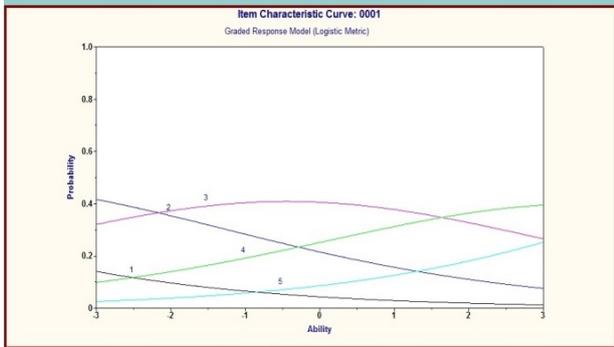


Figura 5: Curva de Informação do Teste

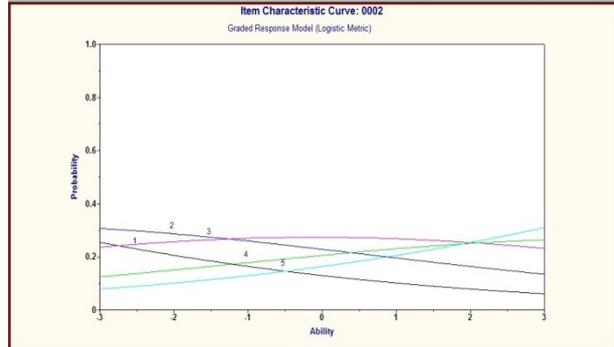
#### 4.2.3 Analisando Item a Item:

Concluindo a análise TRI, são gerados gráficos item a item chamados de Curva Característica do Item mostrando a relação da equação (9) entre os níveis de atitudes dos respondentes (traço latente) e a probabilidade de escolha de cada categoria de resposta. (i) Item 1: a figura 6a mostra que respondentes com traço latente abaixo de -3 (pouco pré dispostos) tem uma probabilidade maior de responder “Não concordo totalmente” representada pela curva 2 e pouca probabilidade de responder “Concordo totalmente” representada pela curva 5. Já para traços latentes maiores (acima de 3) a probabilidade maior é que a resposta ao item seja: “Concordo parcialmente”; (ii) Item2: a figura 6b representa há probabilidade maior de responder não concordo parcialmente para indivíduos não pré dispostos e concordo totalmente para indivíduos com forte traço latente; (iii) Item 3: na figura 6c as curvas demonstram que o item foge do padrão das curvas características para traços latentes, demonstrando aqui uma questão tendenciosa, pois a pré disposição do respondente esta ligada diretamente e proporcionalmente com a escolha da resposta; (iv) Item 4: já a figura 6d demonstra que para grande pré disposição a resposta é concordo totalmente, pouca pré disposição “não concordo parcialmente”. O que demonstra um problema pra empresa, pois só um cliente muito convencido do conhecimento técnico da equipe, marcaria essa opção. Logo indica que este conhecimento não esta sendo demonstrado para os clientes. (v) Item 5: a figura 6e demonstra em suas curvas uma certa tendência a responder “concordo totalmente”, porém essa probabilidade se demonstra alta para traços latentes altos e baixos o que

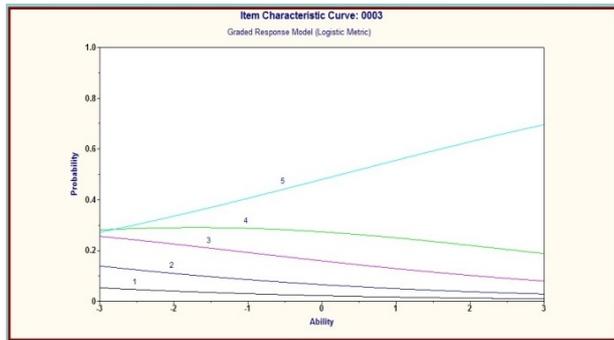
demonstra que a cordialidade dos atendentes é um ponto positivo na empresa. (vi) Item 6: a figura 6f indica que tanto para pouco traço latente quanto para alto a probabilidade de resposta se mantém. Pode indicar um item favorável a empresa.



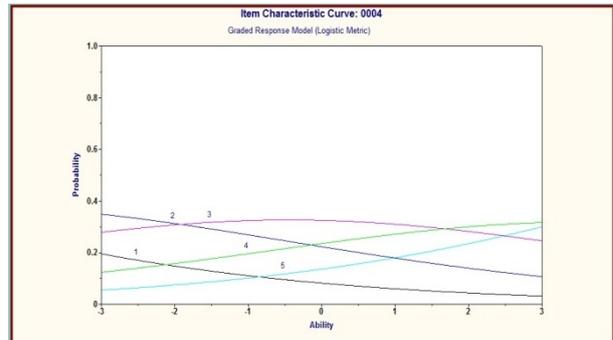
a. Item 1: Quanto ao atendimento, a Comercial superou suas expectativas:



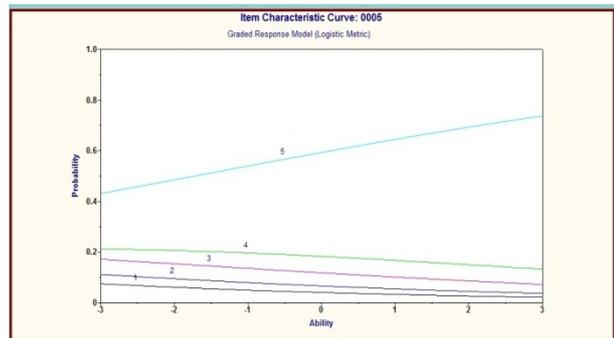
b. Item 2: Quanto as informações, os vendedores as passaram claramente:



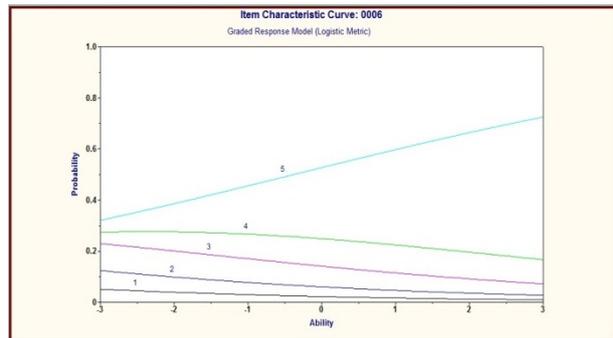
c. Item 3: Quanto a agilidade do atendimento:



d. Item 4: Quanto a equipe, demonstrou-se conhecimento técnico:



e. Item 5: Quanto a cordialidade dos atendentes:



f. : Item 6: Quanto ao suporte pós venda foi satisfatório:

Figura 6: Curvas características ao item

#### 4.2.4 Analisando respondente a respondente:

Para o resultado final da análise, o software Parscale gera os escore da TRI (resultado da equação de Samejima) para cada respondente com base nas respostas dos itens (representada em parte pela tabela 3). Escores negativos indicam respondentes desfavoráveis perante os itens, escores positivos indicam respondentes concordantes com os itens. Com base na análise da frequência, conclui-se que 48% dos clientes estão insatisfeitos e 52% satisfeitos conforme tabela 3.

Cliente	Respostas aos itens						Score TCM (média)	Score TRI (Samejima)
	01	02	03	04	05	06		
0001	3	3	2	3	4	2	2,83	-0,5517
0002	2	4	3	2	2	3	2,67	-0,8171
0003	1	5	4	4	5	1	3,33	-0.2469
0004	1	1	1	3	5	2	2,17	-1.0333
0005	3	4	2	4	5	3	3,50	-0.0998
0006	3	4	3	4	5	5	4,00	0.2364
0007	4	4	2	2	5	2	3,17	-0.2329
0008	4	4	4	5	5	4	4,33	0.4996
0009	4	4	5	5	5	4	4,50	0.6996
0010	4	4	5	4	5	3	4,17	0.4934
...	...	...	...	...	...	...	...	...
0017	5	4	4	3	4	4	4,00	0.2190
0018	4	4	4	3	4	3	3,67	0.0020
0019	5	4	4	4	4	2	3,83	0.2383
0020	5	5	5	4	4	3	4,33	0.5857
0021	4	3	5	3	5	4	4,00	0.3130
0022	3	2	4	4	5	4	3,67	-0.0629
...	...	...	...	...	...	...	...	...
1043	3	5	5	3	5	4	4,17	0,9123
Indicador final: %							=>3 80%	positivos 52%

Tabela 3 – Comparação dos resultados

Estes resultados apontam diferença entre o emprego das duas teorias. Analisando, por exemplo, o cliente 0003, como base no método TCM, a média de resposta é de 3,33, nota acima de 3, o que é considerado um cliente satisfeito com este item. Contudo, analisando pela TRI o mesmo cliente apresenta *score* negativo, indicando um cliente insatisfeito. Isso acontece pelo método de análise diferenciado da TRI que analisa os itens um a um como únicos e depois de forma global, analisando a

postura dos entrevistados um a um diante das questões. O cliente número 0003 pontuou notas baixas para itens no qual o traço latente de resposta afirmativa é alto. Nestes itens há uma tendência a uma resposta afirmativa mascarando uma reação negativa o que levanta a hipótese que este cliente seja no futuro um possível reclamante para o SAC. Pela teoria clássica de medidas, a empresa não consideraria essa a possibilidade.

## 5. Conclusão:

Pesquisas de satisfação são muito difundidas nas empresas como forma de mensurar seu serviço ou produto, e para tanto precisam ser representativas, terem um indicador preciso da real situação das empresas. Na literatura encontra-se variadas formas de pesquisa e análise, porém muitas empresas utilizam o método da teoria clássica de medidas por sua simplicidade na análise de seus resultados. Pesquisando na literatura outros métodos, o método TRI, destaca-se por analisar item a item, respondente a respondente prevendo um resultado mais completo e preciso para pesquisas e testes.

Com a análise feita através da TRI espera-se resultados mais minuciosos e representativos para auxiliar a tomada de decisão da empresa. Aplicou-se o método TRI e TCM na análise do banco de dados da pesquisa da empresa estudada, apontando diferenças entre os valores obtidos. A porcentagem de 52% satisfeitos concluída pela TRI é muito abaixo dos 80% indicados pela TCM, o que faria a empresa agir sobre o serviço de vendas já que identificaria problemas. Como complementaridade, a TRI ainda gera uma análise item a item, deixando de forma explícita a quais itens a empresa necessita concentrar ações de melhoria.

Para o caso estudado, além destes fatores, a TRI indica que a pesquisa tem graves erros, não discrimina seus itens e estes são insuficientes para o resultado ser válido para toda a população de respondentes. Entretanto utilizando somente a TCM a empresa não terá conhecimento deste problema.

A TRI além de analisar a validade da pesquisa é mais criteriosa na avaliação dos itens e análise das respostas dos respondentes. Este artigo apresentou as etapas e discussão de análise pretendendo contribuir para difusão da TRI como alternativa aos instrumentos de avaliação da qualidade em serviços ou produtos.

## REFERÊNCIAS:

ALBUQUERQUE, A.; TROCCOLI, B. T. **Desenvolvimento de uma escala de bem-estar subjetivo**. *Psicologia: Teoria e Pesquisa*, Brasília, v. 20, n. 2, p. 153-164, ago. 2004.

COSTA, M. B. F. **Técnica derivada da Teoria de Resposta ao Item (TRI) aplicada ao setor de serviços**. Dissertação de Mestrado, UFPR, Curitiba, 2001.

DRUCKER, P. F. **The information executives truly need**. *Harvard Business Review*, v.73, n.1, p.54-62, 1995.

EMBRETSON, S.; REISE, S. P. **Item Response Theory for Psychologists**. New Jersey: Lawrence Erlbaum Associates, Inc. Publishers, 2000.

FLETCHER, P.: A teoria da resposta ao item: medidas invariantes do desempenho escolar. **Ensaio: avaliação e políticas públicas em educação**, 1(2), 21-28, 1994.

GUEWEHR, K. **Teoria da resposta ao item na avaliação de qualidade de vida de idosos**. Dissertação de mestrado - UFRGS, Faculdade de Medicina, Programa de Pos-Graduação em Epidemiologia, RS, 2007.

JURAN, J.M.: **Juran na liderança pela qualidade**: um guia para executivos. 2.ed. São Paulo,. Pioneira, 1993.

LAZARFELD, P.F.: **Studies in Social Psychology in World War II**: v. 4. Measurement and Prediction. Princeton, NJ: Princeton University Press. 1950.

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

MÜLLER, Cláudio J.: **Modelo de gestão integrando planejamento estratégico, sistemas de avaliação de desempenho e gerenciamento de processos (MEIO – Modelo de Estratégia, Indicadores e Operações)**. 2003. Tese (Doutorado em Engenharia) – Universidade Federal do Rio Grande do Sul.

NUNES, Carlos Henrique Sancineto da Silva ; PRIMI, Ricardo ; NUNES, Maiana Farias Oliveira ; MUNIZ, Monalisa ; CUNHA, Tatiana Freitas da ; COUTO, Gleiber. **Teoria de Resposta ao Item para Otimização de Escalas tipo Likert: um exemplo de aplicação**. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, v. 25, p. 51-80, 2008.

PASQUALI, L.; PRIMI, R. **Fundamentos da Teoria da Resposta ao Item – TRI**. *Avaliação Psicológica*, v.2 n.2, p. 99-110, 2003.

SAMEJIMA, F. A.: Estimation of latent ability using a response pattern of graded

*scores. Psychometric Monograph*, 1969.

TEZZA, R. ; BORNIA, A. C.:**Teoria da Resposta ao Item: vantagens e oportunidades para a engenharia de produção**. In: XXIX Encontro Nacional de Engenharia de Produção, 2009, Salvador BA. XXIX ENEGEP e XV ICIEOM. Rio de Janeiro - RJ : ABEPRO, 2009.

TURIK, CLAUDIA ,**Análise de atitudes de alunos universitários em relação à estatística por meio da Teoria da Resposta ao Item (TRI)** . Dissertação de mestrado- Programa de Pós Graduação em Educação, PUCRS, Porto Alegre-RS, 2010.

VAN DER LINDEN, W.J. & HAMBLETON, R.K. :**Handbook of modern item response theory**. New York: Springer-Verlag, 1997.

VARGAS, V.C.C., SELIG, P.M., ANDRADE, D.F., RIBEIRO, J.L.D. **Avaliação dos intangíveis: uma aplicação em capital humano**. Gestão e Produção, v. 15, n.3 São Carlos, 2007, p.619-634.

VIANNA, H.M. **Testes em Educação**. São Paulo: Ibrasa. (1987).

WELLIANDRE Alexandre, J.. C.; ANDRADE. D. F.; VASCONCELOS, A. P. & ARAUJO. 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**. Revista Gestão & Produção, 2002.