

---

# O Uso da Técnica de Modelagem em Equações Estruturais na Área de Marketing: um Estudo Comparativo entre Publicações no Brasil e no Exterior

---

Vinícius Andrade Brei  
Guilherme Liberali Neto

## RESUMO

Este artigo realiza uma avaliação crítica da aplicação de Modelagem em Equações Estruturais no Brasil, a partir de um estudo comparativo entre os trabalhos publicados na sub-área de marketing do ENANPAD e nas seções de marketing de três dos principais periódicos de administração do Brasil: RAE, RAUSP e RAC, de 1994 a 2003. Os resultados obtidos foram comparados às estatísticas que resumem a utilização de SEM em artigos publicados nos principais *journals* internacionais de marketing, tabulados e apresentados por Baumgartner e Homburg (1996). O texto se inicia por uma breve visão geral de SEM - surgimento, funções e características, etapas principais, aplicações, aspectos essenciais e alguns aspectos problemáticos da sua utilização. Após a discussão sobre os critérios utilizados para tabular os artigos, o panorama do uso de SEM no Brasil é apresentado e comparado com as estatísticas dos trabalhos publicados em âmbito internacional. O texto é encerrado com uma avaliação geral da qualidade da utilização desta abordagem no Brasil e apresentação de sugestões para futuras aplicações da modelagem de equações estruturais em marketing.

**Palavras-chave:** modelagem em equações estruturais; marketing; métodos multi-variados.

## ABSTRACT

This article investigates the use of Structural Equation Modeling (SEM) in the published articles on the main brazilian business administration conference (ENANPAD) marketing track and also on the marketing section of the most important brazilian business administration journals (RAE, RAUSP and RAC) in the last ten years: 1994 to 2003. In addition to the analysis of the published articles, the authors present a comparison with Baumgartner and Homburg's (1996) findings in their investigation on SEM usage, which explored the papers published at main international marketing journals. The text begins with a brief overview of SEM – its origins, functions and characteristics, main steps, some applications, its essentials and some problematic aspects of the method. Then the criteria used to analyze the papers are presented. The results are shown in detail, including a comparison with Baumgartner and Homnburg's findings. This article ends with an overall evaluation about the quality of the use of SEM in Brazil, along with some suggestions for future applications of SEM in marketing.

**Key words:** structural equation modeling; marketing; multivariate methods.

## INTRODUÇÃO

O uso da Modelagem em Equações Estruturais (daqui para frente chamada pela abreviação do nome original em inglês – *Structural Equation Modeling* – SEM) recebeu grande impulso por volta do final da década de 70 e início dos anos 80 do século passado, com o trabalho de Bentler (1980) e com a popularização do *software* LISREL (Steenkamp & Trijp, 1991). Desde então, tem sido cada vez mais utilizada dentro do campo das pesquisas sociais e comportamentais (Baumgartner & Homburg, 1996).

Ao trabalho de Bagozzi (1980) costuma ser creditada a primeira aplicação da técnica na área de marketing e, mais especificamente, nas pesquisas sobre comportamento do consumidor. SEM vem tendo, desde então, utilização crescente, como pode ser observado em muitos periódicos internacionais de marketing.

Em trabalho de revisão da aplicação de SEM em marketing, Hulland et al. (1996) e Baumgartner e Homburg (1996) identificaram 186 publicações entre 1980 e 1994, e 149 publicações entre 1977 e 1994, respectivamente. Apenas para efeito de exemplificação da sua crescente popularidade, MacCallum e Austin (2000) analisaram aproximadamente 500 publicações em que a técnica SEM foi aplicada, num total de 16 *journals* de psicologia entre 1993 e 1997. James e James (1989) identificaram 55 aplicações na área de pesquisa organizacional, tendo Medsker, Williams e Holohan (1994) identificado outras 28.

No Brasil, a modelagem em equações estruturais demorou a ser utilizada na área de marketing, em comparação ao cenário internacional anteriormente descrito. Por exemplo, considerando o período de dez anos (1994-2003), a primeira utilização de SEM em marketing no ENANPAD – Encontro Nacional do Programa de Pós-Graduação em Administração, usualmente o fórum de ponta em termos de pesquisas na área de administração no País, data apenas de 1998, tendo, porém, aumentado nos anos seguintes.

A crescente popularidade de SEM no exterior e no Brasil pode ser creditada a dois fatores. Em primeiro lugar houve, nos últimos anos, grande popularização de programas de computador para utilização de SEM, como LISREL, EQS, AMOS, COSAN, EZPATH, LINCS, PROC CALIS (do SAS), RAMONA, e outros. Além disso - e mais importante - a abordagem SEM tem bom potencial para investigações empíricas abrangentes de aspectos teóricos e de mensuração

(Baumgartner & Homburg, 1996). Se por um lado, esses fatores favorecem a utilização da abordagem, por outro eles são motivo de preocupação sobre eventuais usos inadequados ou problemáticos de SEM em pesquisas empíricas. A preocupação é bastante pertinente, na medida em que existem muitos aspectos nos diversos estágios de aplicação da mesma que, se não forem observados, podem gerar não só resultados errôneos, mas também conclusões inadequadas ou imprecisas.

A partir dessas observações, este artigo realiza uma avaliação crítica da aplicação de SEM no Brasil, a partir de um estudo comparativo entre os trabalhos publicados na subárea de marketing do ENANPAD e nas seções de marketing de três dos principais periódicos de administração do Brasil: RAE, RAUSP e RAC. O período analisado, em ambos os casos, foi de 1994 a 2003 (entretanto, a primeira edição da RAC data de janeiro de 1997). Os resultados obtidos foram comparados às estatísticas que resumem a utilização de SEM em artigos publicados nos principais *journals* internacionais de marketing, tabulados e apresentados por Baumgartner e Homburg (1996). Além disso, para efeito de enriquecimento das comparações, também foram consultados o trabalho de Jarvis, Mackenzie e Podsakoff (2003) e a revisão de MacCallum e Austin (2000).

## VISÃO GERAL DA MODELAGEM EM EQUAÇÕES ESTRUTURAIS

O termo Modelagem em Equações Estruturais não designa apenas uma única técnica estatística, mas uma família de procedimentos relacionados (Kline, 1998). Pode-se considerar que sua função principal é a especificação e estimação de modelos de relações lineares entre variáveis. Essas variáveis podem incluir tanto variáveis mensuráveis (VM) quanto variáveis latentes (VL). VL's são construtos hipotéticos que não podem ser mensurados diretamente. Em SEM, tais construtos são tipicamente representados por múltiplas VM's que servem como indicadores dos construtos. Assim, um modelo de equações estruturais é um suposto padrão de relacionamentos lineares entre um conjunto de VM's e VL's. Apesar do termo SEM designar um conjunto de técnicas ou métodos, algumas funções e características em comum podem ser identificadas entre elas, tais como as descritas a seguir (MacCallum & Austin, 2000).

### Funções e Características

Em primeiro lugar, SEM é determinada *a priori*; ou seja, os programas de computador que automatizam alguns dos procedimentos estatísticos de SEM

requerem que o pesquisador forneça diversas informações como, por exemplo, quais variáveis afetam o quê, e quais as direções desses efeitos. Esses relacionamentos entre variáveis representam as hipóteses do pesquisador. Neste sentido, SEM é vista como uma técnica confirmatória: o modelo deve ser desenvolvido antes da coleta dos dados, e o principal questionamento é se os dados o confirmam ou não.

Entretanto este suporte empírico pode, por vezes, não ser encontrado. Nestes casos, o pesquisador tem a opção de abandonar o modelo ou modificar as hipóteses em que ele é baseado. A primeira opção é a mais drástica e a menos usual. A segunda costuma ser mais utilizada, e é quando a técnica SEM adquire uma tendência um tanto quanto **exploratória**, na medida em que modelos revisados são testados no mesmo conjunto de dados empíricos.

Joreskog e Sorbom (1993) formalizaram essas idéias, distinguindo três possíveis aplicações de SEM: 1) estritamente confirmatória; 2) utilização de modelos alternativos e 3) geração de modelos. Na primeira aplicação, o pesquisador tem um único modelo que é aceito ou rejeitado, baseado em sua correspondência ou não com os dados. O segundo caso é um pouco mais freqüente que o primeiro, mas ainda restrito a situações em que há mais de um modelo apriorístico. A terceira situação parece ser a mais comum, e ocorre quando o modelo inicial não se ajusta aos dados. Nestes casos, o modelo é modificado pelo pesquisador e é novamente analisado com os mesmos dados, com o objetivo de **descobrir** um modelo que tenha sentido em termos teóricos e que apresente correspondência estatística razoável com os dados coletados.

A estatística básica utilizada para a aplicação de SEM é a matriz de covariância; sendo ainda possível a utilização da matriz de correlações. A vantagem da utilização da covariância como estatística básica é que ela traz mais informação do que um dado padronizado, como a correlação. Em outras palavras, os dois objetivos das análises são entender padrões de correlações entre os conjuntos de variáveis, e explicar a maior parte da variância possível, através do modelo especificado pelo pesquisador. Em termos de utilização, SEM pode ser aplicada para dados não experimentais (correlacionais), experimentais, ou até mesmo uma mistura de ambos.

Importante ressaltar que diversos procedimentos estatísticos como regressão múltipla, correlação canônica, análise fatorial, ANOVA e MANOVA podem ser vistos como casos especiais de SEM. Em outras palavras, a **família** SEM abarca diversas técnicas que têm como característica em comum um modelo linear geral de relacionamento entre variáveis. Essa característica explica a crescente popularidade, robustez e flexibilidade de SEM, na medida em que pode ser vista como um conjunto de técnicas ou procedimentos estatísticos.

Usualmente, SEM requer amostras relativamente grandes, em função dos diversos pré-requisitos inerentes aos tipos de técnicas de análise e procedimentos estatísticos geralmente empregados. Regras empíricas geralmente citam que este número deve variar entre 5 e 10 casos por parâmetro estimado (Bentler & Chou, 1987). Por último, é possível testar muitos tipos de significância estatística em SEM, mas o papel desses testes, na análise como um todo, é menos importante do que em técnicas mais comuns, como na ANOVA. A razão dessa menor importância é que, em SEM, a avaliação de modelos se dá numa perspectiva em nível macro, e não micro. Além disso, testes de significância podem ser influenciados pelo tamanho da amostra, e como anteriormente descrito, SEM costuma requerer amostras relativamente grandes. Assim, alguns desses testes estatísticos podem apresentar significância estatística em função do tamanho da amostra, o que não ocorreria se a mesma fosse relativamente pequena.

Todas as características acima descritas servem como pano de fundo para a aplicação de SEM, o que usualmente deve ser feita em diversos estágios (ou etapas) sequenciais. A seção a seguir apresenta os elementos fundamentais de cada um.

## Etapas de aplicação de SEM

Há alguma variação entre os autores a respeito da indicação da seqüência dos estágios de aplicação de SEM. Entretanto, em geral, são seguidas as etapas descritas por Hair et al. (1998).

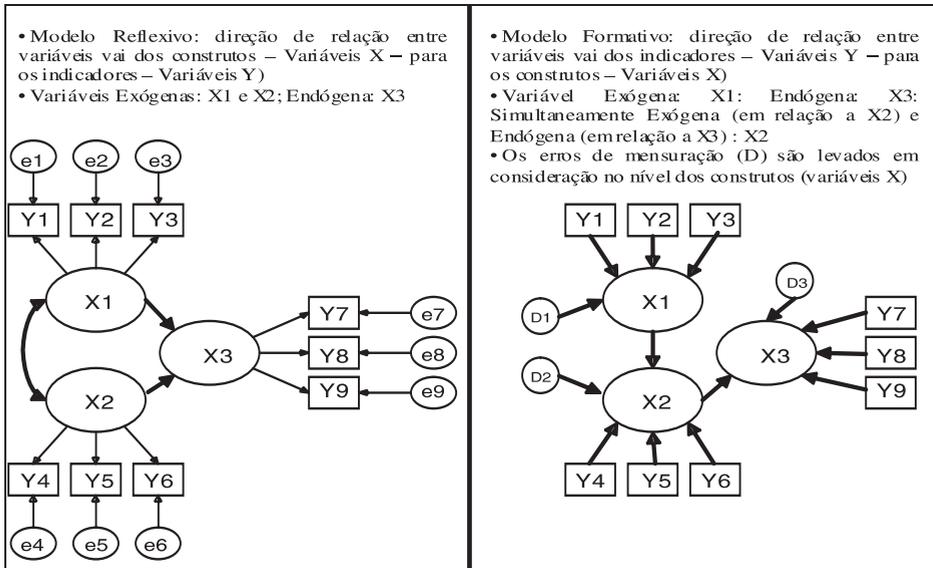
**Etapas 1 e 2 – Desenvolvimento de um Modelo Baseado na Teoria e Construção de um Diagrama de Caminhos de Relações Causais.** SEM é baseada em relações que sugerem evidências de causalidade. Entretanto tais evidências não podem ser efetivamente comprovadas senão através de um *design* de pesquisa criado especificamente para este fim, como seria possível observar, por exemplo, em alguns tipos de estudos verdadeiramente experimentais (em oposição aos quase-experimentais). Entretanto é importante notar que os fatores que permitem investigar a causalidade *stricto sensu* de uma relação entre duas variáveis não se restringem apenas a questões técnicas de análise de dados. Eles incluem uma série de aspectos epistemológicos, teóricos e metodológicos, cuja discussão foge do escopo deste trabalho. Ainda assim, especificamente em relação à análise de dados, há alguns critérios usualmente utilizados para a investigação de evidências de causalidade. Hair, Anderson, Tatham, & Black (1998), por exemplo, citam os seguintes: 1) associação suficiente entre duas variáveis; 2) antecedência temporal da causa sobre o efeito; 3) ausência de variáveis causais alternativas (o chamado **erro de especificação**); e 4) base teórica para o relacionamento.

A visualização gráfica dessas relações é o que se denomina **diagrama de caminhos**, utilizado não apenas para a identificação de relações de causa e efeito entre os construtos (relacionamentos entre variáveis dependentes e independentes), mas também para relacionamentos derivados (correlações) entre construtos e até mesmo indicadores. Portanto os elementos básicos de um diagrama de caminhos são o construto e setas: uma seta reta representa relação causal direta de um construto em relação a outro, enquanto uma seta curva entre quaisquer variáveis indica uma correlação entre elas. As variáveis independentes do modelo são chamadas **exógenas**, pois não são **causadas** ou **preditas** por nenhuma outra variável no modelo (uma indicação clara de que uma variável é exógena é a ausência de setas apontadas para ela). Por fim, as variáveis **endógenas** ou dependentes são preditas por outros construtos exógenos (vide Figura 1).

**Etapa 3 – Conversão do Diagrama de Caminhos para um conjunto de Modelos de Mensuração e Estrutural.** Seguindo as orientações de Hair et al. (1998) e de Anderson e Gerbin (1988), o pesquisador deve usar SEM em dois passos distintos, quando quer obter uma representação acurada da confiabilidade dos indicadores utilizados. Assim, primeiramente estima-se o modelo de mensuração e, posteriormente, estima-se o modelo estrutural. A finalidade principal do modelo de mensuração é verificar se os itens operacionais utilizados para medir os construtos são significativos e realmente medem aquilo que se esperava que medissem (isto é, verificação da validade de construto). Assim, ele permite avaliar se cada construto tem confiabilidade satisfatória para se estimarem as relações causais em etapas posteriores de teste do modelo.

Nesta etapa do desenvolvimento da modelagem, é essencial que o pesquisador analise as características do fenômeno em estudo, e decida se o modelo apresenta características formativas ou reflexivas. Num modelo reflexivo, a direção de causalidade vai do construto para seus indicadores. Em outras palavras, neste tipo de modelo, mudanças no construto causam mudanças nos itens. Já num modelo formativo, é esperado que variações nos itens causem mudanças no construto ao qual eles estão vinculados (vide Figura 1). As implicações de erros de enquadramento são severas: tratar um modelo formativo como reflexivo implica analisar os antecedentes de um construto como se fossem seus itens, o que compromete a robustez do modelo e das conclusões obtidas a partir dele (Jarvis et al., 2003).

**Figura 1: Exemplos de Modelos Reflexivos/Formativos e de Variáveis Endógenas/Exógenas**



Para testar se os itens operacionais utilizados para medir os construtos atingiram níveis aceitáveis de confiabilidade e de validade lógica, de estrutura interna e nomológica, podem ser realizadas Análises Fatoriais Confirmatórias (AFC) (Byrne, 1994). A AFC pode ser vista como uma das possíveis utilizações da Modelagem em Equações Estruturais (Pedahzur & Schmelkin, 1991), e é desenvolvida através da especificação de modelos com várias finalidades como, por exemplo, desenhos para validação de construtos, refinamento de escalas e mensuração de invariância (MacCallum & Austin, 2000). Mais especificamente, o desenvolvimento de medidas para construto, usando a AFC, pode dar-se através do teste em um modelo de mensuração das supostas relações entre os indicadores (variáveis observáveis) e seus fatores/construtos (variáveis latentes), bem como das relações entre estes últimos (Pedahzur & Schmelkin, 1991). Depois de definido o modelo de mensuração, estima-se o modelo estrutural através dos parâmetros ou coeficientes do diagrama de caminhos. Esta estimação é resultado da conversão desse diagrama em um conjunto ou sistema de equações, que representam matematicamente a relação gráfica (ou estrutural) entre as variáveis.

**Etapa 4 - Escolha do Tipo de Matriz para Entrada de Dados e Estimação do Modelo Estrutural.** A técnica SEM difere de outras técnicas multivariadas, pois utiliza apenas matrizes de variância/covariância ou correlação como entrada de dados. Mesmo que observações individuais sejam inseridas, elas são convertidas em um desses dois tipos de matrizes. Isto ocorre porque o foco de SEM não é em observações

individuais, mas sim no padrão de relacionamentos obtido a partir das respostas dos indivíduos (Hair et al., 1998). Conforme explicado anteriormente, a matriz de covariância traz mais informação do que um dado padronizado, como a correlação.

Esta etapa também prevê a escolha de qual a técnica de estimação deverá ser usada no cálculo do modelo estrutural. Em marketing, observa-se que se costuma utilizar a *Maximum Likelihood* (ML) - padrão na maioria dos programas – a qual gera parâmetros mais robustos, quando o pressuposto de normalidade dos dados é respeitado. Caso haja problemas de normalidade, outros métodos podem ser cogitados, como, por exemplo, o Mínimos Quadrados (Gujarati, 2000) ou o *Elliptical Reweighted Least Squares* (ERLS).

**Etapa 5 – Verificação da Identificação do Modelo Estrutural.** Em termos simples, um problema de identificação é a inabilidade do modelo proposto em gerar estatísticas únicas, isto é, significativas e lógicas (Hair et al., 1998). Uma regra trivial, porém eficaz, é que se deve ter mais parâmetros conhecidos do que parâmetros livres, o que gera o chamado modelo sobreidentificado (*overidentified model*).

**Etapa 6 – Avaliação de Critérios de Ajuste do Modelo.** Existem três grandes grupos de medidas de ajuste: as absolutas, as incrementais e as de parcimônia. As medidas de adequação absolutas determinam o grau em que o modelo geral (tanto o estrutural, quanto o de mensuração) prediz a matriz de covariância ou de correlação. As medidas de ajuste incrementais comparam o modelo proposto a algum modelo básico, chamado de modelo nulo, aquele em que um único construto, com seus indicadores, mede perfeitamente o referido construto. Em outras palavras, representa o valor da estatística qui-quadrado associado à variância total em um grupo de correlações e de covariâncias (Hair et al., 1998, p. 657). As medidas de ajuste de parcimônia relacionam o ajuste do modelo ao número de coeficientes estimados requeridos para atingir o nível de adequação obtido. O objetivo básico é diagnosticar se o ajuste do modelo foi atingido pelo **sobreajuste** (*overfitting*) dos dados com o uso de muitos coeficientes.

**Etapa 7 – Interpretação e Modificação do Modelo.** Uma vez considerados adequados em termos de ajuste, os resultados obtidos devem ser confrontados com a teoria proposta, para que algumas perguntas possam ser respondidas (Hair et al., 1998): Os relacionamentos da teoria foram comprovados estatisticamente e substancialmente? Os modelos concorrentes ensejam a formulação de hipóteses alternativas? Os relacionamentos entre as variáveis ocorreram na direção proposta pelas hipóteses?

Além disso, outro aspecto a ser levado aqui em consideração é o uso de soluções padronizadas ou não padronizadas. Soluções padronizadas variam de 0 a 1 e suas conclusões são específicas para a amostra em questão, uma vez que não mantêm as propriedades da escala utilizada. Entretanto são de fácil entendimento.

Uma última questão importante é com relação à reespecificação do modelo. Do ponto de vista operacional, ela poderia ser feita continuamente, dado que os programas usados sugerem inserção ou cancelamento de variáveis que melhorariam os índices de ajuste do modelo. Ainda assim, toda e qualquer modificação no modelo só deve ser feita com base na teoria, a não ser que alguma conclusão empírica suporte, fortemente, o levantamento de novas hipóteses que questionem a teoria existente (Byrne, 1994).

Descritas as funções, as características e os estágios de aplicação de SEM, serão apresentados, a seguir, os critérios de análise utilizados nesta pesquisa, além do resultado da avaliação dos artigos.

## APLICAÇÕES DE SEM EM MARKETING NO BRASIL

Do ano 1994 até 2003, foram identificados 36 artigos que utilizaram SEM, que são a base para as análises descritas abaixo. Deste grupo, seis foram publicados nos periódicos RAE, RAUSP ou RAC, e 30 foram apresentados e publicados no ENANPAD (a lista completa dos artigos analisados pode ser obtida mediante contato com os autores). É interessante notar que 64% destes artigos lidam com construtos específicos da subárea de comportamento do consumidor, como: satisfação, lealdade, confiança, e outros. Os artigos restantes (36%) ficaram espalhados em orientação para mercado, marketing de relacionamento, serviços e outros. Tal fenômeno pode ser considerado como um resultado não surpreendente, já que as raízes do uso de SEM remontam justamente à área de psicologia, como uma técnica para mensuração psicométrica.

### Critérios de Avaliação

Os trabalhos foram analisados a partir dos critérios sugeridos por Bagozzi e Baumgartner (1994), e utilizados em Baumgartner e Homburg (1996) para a revisão da aplicação de SEM nos principais periódicos internacionais de marketing, entre 1977 e 1994. Além deste trabalho na área de marketing, também foram utilizados os elementos apontados na revisão de MacCallum e Austin (2000), na área de psicologia, e o artigo de Jarvis et al. (2003) sobre uso de modelagem em pesquisas de marketing e de comportamento do consumidor. Os critérios dividem-se em dois blocos: os ligados à **especificação do modelo teórico e design da pesquisa**, e os ligados à **estimação e teste dos modelos**. Ambos são detalhados a seguir.

Com relação à **especificação dos modelos e do design de pesquisa**, o

primeiro critério é o tipo de modelo, que foi categorizado em 1, 2 ou 3, correspondendo, respectivamente, aos modelos de mensuração confirmatório, aos modelos estruturais com um único indicador e aos modelos integrados de mensuração e estruturais. O segundo critério avaliou se os modelos utilizados pelos artigos foram adequadamente classificados e analisados como formativos ou reflexivos, à luz do trabalho de Jarvis et al. (2003), discutido anteriormente. Utilizou-se, também, um grupo de critérios para avaliar a dimensionalidade e robustez dos construtos: a quantidade de variáveis observadas e latentes, a presença de construtos com apenas um indicador, a presença de erros correlacionados, e a existência ou não de justificativa para os casos em que houve tais correlações. Além disso, foi utilizado um grupo de indicadores para avaliar a relação entre o tamanho da amostra e a quantidade de parâmetros a serem analisados, bem como o atendimento aos valores críticos de número de casos por parâmetro sugerido pela literatura (5 e 10). Por fim, o último grupo de indicadores desse bloco avaliou se os artigos reportaram a identificação de modelos, e a quantidade de graus de liberdade.

A avaliação dos aspectos ligados à **estimação e teste dos modelos** em dados empíricos incluiu indicadores sobre a técnica de estimação utilizada (por exemplo, verossimilhança, mínimos quadrados ou outras) e problemas encontrados na estimação, como os decorrentes de não convergência e valores impossíveis (como variância negativa, por exemplo). Além disso, foram utilizados indicadores para avaliar a frequência e a magnitude média de índices de ajuste absolutos, incrementais e de parcimônia. Foram também utilizados indicadores para mapear o fornecimento de informações sobre a confiabilidade e o poder explicativo dos modelos utilizados nos artigos. Por fim, analisou-se o grau em que os artigos documentaram e fundamentaram reespecificações do modelo original e dos modelos rivais.

Com relação aos indicadores utilizados e relatados em artigos, é necessário citar uma ressalva importante: neste trabalho foram analisados apenas os artigos publicados em revistas ou os apresentados no ENANPAD, e não as teses ou as dissertações que, eventualmente, tenham dado origem a vários desses textos. Considerando o limitado espaço disponível em um artigo (em torno de 16 páginas), é razoável supor que, eventualmente, alguns autores tenham utilizado indicadores adicionais para suas análises, os quais podem ter sido reportados apenas nas teses e dissertações, e não nos artigos. Entretanto, como esta limitação também precisa ser considerada na análise do estudo de Baumgartner e Homburg (1996), é razoável esperar que a comparação entre os resultados deste trabalho e do estudo de Baumgartner e Homburg não tenha ficado prejudicada.

Assim, esses indicadores formam uma grade de análise que permite avaliar os

artigos à luz de todos os estágios de aplicação de SEM, descritos por Hair et al. (1998). Além disso, ela também permite comparar a situação do uso de SEM no Brasil com os resultados encontrados no trabalho de Baumgartner e Homburg (1996) sobre o uso de SEM no exterior, a ser apresentado ao longo da próxima seção.

## RESULTADOS E DISCUSSÃO

No período de dez anos de publicações e ENANPADs, foram encontrados 36 trabalhos envolvendo modelagem em equações estruturais, sendo que um foi publicado em 1998, quatro em 1999, dois em 2000, seis em 2001, nove em 2002 e catorze em 2003. Os resultados desta análise estão na segunda coluna da tabela a seguir. Para efeito de comparação, a terceira coluna da tabela apresenta os resultados obtidos por Baumgartner e Homburg (1996), no estudo sobre artigos publicados nos seguintes periódicos, entre 1977 e 1994: *Journal of Marketing*, *Journal of Marketing Research*, *Journal of Consumer Behavior* e *International Journal of Research in Marketing* (Baumgartner & Homburg, 1996).

**Tabela 1: Aplicações de SEM em Marketing no Brasil e no Exterior – Classificação por Critério de Avaliação**

Critérios	Resultados Consolidados Brasil	Resultados Consolidados Exterior
<b>1. CRITÉRIOS RELACIONADOS À ESPECIFICAÇÃO DOS MODELOS E DESIGN DE PESQUISA</b>		
Tipo de modelo	(Tipo 1=11%; 2=8%; 3=81%)	(Tipo 1=44%; 2=16%; 3=40%)
Modelos corretamente especificados como Formativos ou Reflexivos	81%	71%
Número médio de variáveis observadas	29	11
Número médio de variáveis latentes	6	5
Razão entre variáveis observadas e variáveis latentes (critério <sup>1</sup> : $\geq 3$ )	5,3	2,2
Modelos contendo pelo menos um construto com apenas um indicador	28%	48%
Modelos de mensuração com correlação entre erros de mensuração	3%	10%
Justificativa para correlação entre erros de mensuração (critério: 100%)	0%	53%
Modelos não recursivos	NI	17%
Modelos estruturais com correlação entre erros	NI	14%
Justificativa para correlação entre erros nos parâmetros estruturais (critério: 100%)	NI	27%
Tamanho médio de amostra	801	178
Número de Parâmetros Estimados <sup>2</sup>	64	29
Razão entre tamanho da amostra/ n.parâmetros (critério mínimo: $\geq 5$ ; desejável $\geq 10$ )	13	6,4
Modelos com razão n/parâmetros livres $<5$	31%	41%
Modelos com razão n/parâmetros livres $<10$	64%	73%
Modelos cujas evidências de identificação foram apresentadas (critério: 100%)	NI	5%
Número médio de graus de liberdade	142	32

(continuação)

**Tabela 1: Aplicações de SEM em Marketing no Brasil e no Exterior – Classificação por Critério de Avaliação**

Critérios	Resultados Consolidados Brasil	Resultados Consolidados Exterior
<b>2. CRITÉRIOS RELACIONADOS À ESTIMAÇÃO E TESTE DOS MODELOS</b>		
Modelos estimados por ML	61%	95%
Modelos onde houve problemas de convergência	NI	1%
Modelos com solução inadequada (critério: 0%)	NI	5%
Modelos onde pelo menos um índice stand-alone de ajuste foi reportado (critério: 100%)	86%	92%
Frequência de uso do teste Qui-quadrado	92%	90%
Frequência de uso do GFI	72%	46%
Frequência de uso do AGFI	69%	28%
Frequência de uso do RMR	22%	39%
Frequência de uso de outros índices stand-alone	86%	10%
Tamanho médio da razão Qui-quadrado /GL (critério: $\leq 5$ )	6,81	1,62
Magnitude média do GFI (critério: o mais próximo possível de 1)	0,88	0,95
Magnitude média do AGFI (critério: $\geq 0,90$ )	0,76	0,91
Magnitude média do RMR (critério: $\leq 0,05$ )	0,23	0,05
Magnitude média do RMSEA (critério: $\leq 0,08$ )	0,09	0,06
Modelos onde pelo menos um índice incremental de ajuste foi reportado (critério: 100%)	92%	38%
Frequência de uso do BBI (%)	0%	23%
Frequência de uso do TLI (%)	42%	17%
Frequência de uso do CFI (%)	69%	13%
Frequência de uso outros índices incrementais (%)	72%	3%
Magnitude média do BBI (critério: $\geq 0,90$ )	N.A.	0,91
Magnitude média do TLI (critério: $\geq 0,90$ )	0,84	0,93
Magnitude média do CFI (critério: $\geq 0,90$ )	0,92	0,95
Magnitude média do RFI (critério: $\geq 0,90$ )	0,88	0,85
Magnitude média do IFI (critério: $\geq 0,90$ )	0,90	0,95
Modelos para os quais algum tipo de informação sobre confiabilidade dos modelos foi reportada (critério: 100%)	75%	78%
Modelos para os quais o $R^2$ foi reportado (critério: 100%, quando há modelos estruturais)	33%	23%
Modelos para os quais foi realizada uma discussão sobre reespecificações de modelo (critério: 100%, quando houve reespecificação)	36%	54%
Modelos para os quais foi realizada comparação de modelos	36%	31%
Estudos para os quais os modelos alternativos foram sustentados conceitualmente ( <i>Cross-validated</i> )	31%	21%

Fonte: Anais do ENANPAD, RAE e RAUSP (1994-2003), RAC (1997-2003) e adaptação de Baumgartner e Homburg (1996).

Notas: NI = Não Informado. NA = Não Aplicável.

<sup>1</sup> Nos casos onde há definição de critérios/parâmetros usualmente aceitos na literatura, os mesmos foram incluídos. Fontes: Baumgartner e Homburg (1996), Bentler e Chou (1987), Bollen (1989) e DeVellis (1991).

<sup>2</sup> Nos vários casos em que os autores não informaram adequadamente o número de parâmetros estimados, decidiu-se estimar o valor mínimo de parâmetros estimáveis de acordo com o seguinte procedimento: dois parâmetros por item (referentes à carga fatorial ou coeficiente de regressão e o erro) mais o número de variáveis latentes.

## **Análise dos Critérios Relacionados à Especificação dos Modelos e do Design de Pesquisa**

Comparando-se os trabalhos publicados no Brasil (em revistas e nos Encontros da ANPAD) com os do exterior, observa-se a predominância, no Brasil, de pesquisas que utilizam modelos Tipo 3 (81% dos casos, contra 40% no exterior), geralmente os mais complexos e robustos, por integrarem modelos de mensuração com modelos estruturais. Talvez este ponto possa ser explicado, em parte, pela diferença no *timing* da análise: os artigos nacionais são resultados de pesquisas conduzidas mais recentemente do que os estrangeiros. Outro aspecto positivo identificado nas pesquisas brasileiras é o menor percentual de modelos Tipo 2 (modelos estruturais com um único indicador): 8% contra 16%. Este tipo de modelo pode apresentar problemas relacionados à mensuração da confiabilidade, dado que cada construto é medido por apenas uma variável (Baumgartner & Homburg, 1996).

O total de modelos que contêm ao menos uma variável latente (i.e. com apenas um indicador) foi inferior no Brasil, em comparação às publicações internacionais (28%, contra 48%), demonstrando uma maior preocupação dos autores que publicaram trabalhos no País em relação à robustez e dimensionalidade dos construtos nas pesquisas em que SEM foi usada. Isto é corroborado pela razão entre o número de variáveis observadas e o número de variáveis latentes: no Brasil (doravante, apenas BR), obteve-se 5,3 e no exterior (doravante INT'L) Baumgartner e Homburg obtiveram 2,2. Aqui é importante lembrar a sugestão de Bollen (1989): cada variável latente deve ser medida por um mínimo de três ou quatro indicadores (DeVellis, 1991).

Portanto, em termos de quantidade de itens, as pesquisas brasileiras estão utilizando adequadamente a relação indicada para aplicação de SEM. Entretanto esta constatação positiva precisa ser amenizada pelo fato de que não foi possível avaliar a unidimensionalidade de forma completa e precisa: vários autores não reportaram erros de mensuração correlacionados, nem a eventual inclusão de itens com cargas cruzadas (isto é, um mesmo item servindo de indicador de duas variáveis latentes distintas). Recomenda-se que a existência ou inexistência de correlações ou de cargas cruzadas seja explicitamente relatada, pois a completa e correta unidimensionalidade de construtos é uma característica desejável dos modelos de mensuração (Baumgartner & Homburg, 1996, p. 146).

Um comentário negativo em relação aos trabalhos nacionais é que, nos estudos onde foi reportada correlação entre erros, não foram apresentadas justificativas para a existência de tal correlação (BR=0%, INT'L=53%). Uma possível hipótese para essa não exploração das causas é que, dada a recentidade da aplicação de SEM no País, os autores – talvez por questão de espaço - e os revisores ainda não estejam atentos a detalhes como esses. Entretanto, esses fatores podem gerar estimativas de parâmetros de maior magnitude, e significâncias estatísticas

inexistentes. Além disso, a existência de correlação entre erros pode estar indicando que há construtos que não foram incluídos no modelo, mas que podem estar influenciando o comportamento do fenômeno estudado e, portanto, mereceriam investigação teórica adicional.

O tamanho da amostra é um dos fatores mais relevantes, quando se utiliza SEM, dado que análises realizadas em pequenas amostras frequentemente são problemáticas (MacCallum & Austin, 2000, p. 216). Não há uniformidade na literatura em relação ao número absoluto do que pode ser considerada uma amostra pequena (podem ser encontrados valores entre 100 e 200 casos). Entretanto há muito mais acordo no que diz respeito à razão entre tamanho da amostra/número de parâmetros estimados no modelo. Bentler e Chou (1987) defendem que esta razão deve ser de pelo menos 5:1, para que se consigam estimativas consistentes de parâmetros. Além disso, segundo esses autores, esta relação deve ser ainda maior (pelo menos 10:1), para que se obtenham testes de significância estatística apropriados.

A análise comparativa, tanto em relação ao tamanho médio da amostra (BR=801; INT'L=178) como à razão entre tamanho da amostra/número de parâmetros (BR=13; INT'L=6,4) foram mais positivas no Brasil do que no exterior. Ainda que o tamanho médio de amostra no Brasil tenha sido muito influenciado por um dos trabalhos (cuja amostra foi de 10935 casos), a sua eliminação reduz o tamanho médio de amostra para 449 casos, o qual ainda é mais de 2,5 vezes maior que a média internacional. A eliminação de *outliers* foi realizada apenas no critério **tamanho médio da amostra**, porque, no Brasil, foi encontrado apenas um trabalho com tamanho de amostra extremamente influenciador da média, o que não ocorreu no caso das outras medidas. Como no estudo internacional não houve eliminação de outliers, decidiu-se seguir o mesmo critério para evitar viés nos resultados comparativos.

No caso brasileiro, cabem ainda duas ressalvas. A primeira é positiva: a média de 13 casos por parâmetro estimado ficou acima do patamar sugerido na literatura. A ressalva negativa diz respeito à falta de informações nos artigos nacionais em relação ao número de parâmetros estimados (cerca de metade dos trabalhos do ENANPAD não trouxeram esta informação), o que dificulta substancialmente a avaliação da aplicação de SEM.

Em números absolutos, os graus de liberdade nos trabalhos brasileiros foram quase 5 vezes maiores que no exterior (BR=142; INT'L=32). Entretanto uma razão bastante diferente foi identificada, quando se analisam os parâmetros estimados, os geradores dos graus de liberdade (BR=64 contra INT'L=29). Esta diferença pode ser um indicativo de que os graus de liberdade (ou os parâmetros estimados) estejam sendo informados de maneira incorreta em muitos estudos nacionais. A possível razão para esse erro ocorre, quando a matriz de correlação é usada como entrada de dados, e pelo ao menos um construto exógeno é mensurado por um único indicador. A carga fatorial desse indicador único em

relação ao seu construto é constricta a um e, ao mesmo tempo, a variância do indicador único é assumida como sabida e fixa em um. O resultado disso é que ocorre uma inflação de graus de liberdade e valores-p exagerados para a estatística qui-quadrado. Nesses casos, o recomendado é que a escala de uma variável latente deve ser fixa, mantendo a carga fatorial do indicador ou a variância do fator (construto) igual a um, mas nunca os dois simultaneamente (Baumgartner & Homburg, 1996, p. 147).

## **Análise dos Assuntos Relacionados à Estimação e Teste dos Modelos Teóricos em Dados Empíricos**

Uma dos principais pressupostos de SEM é a normalidade dos dados. Sendo identificada essa normalidade, a utilização do procedimento de estimação de parâmetros Máxima Verossimilhança (*Maximum Likelihood* – ML) pode ser feita sem problemas. Caso contrário, outros métodos devem ser considerados. ML é o padrão da maioria das implementações SEM em software, o que possivelmente é a explicação para o elevado percentual de utilização no exterior (95%), segundo Baumgartner e Homburg (1996). Nas pesquisas brasileiras, menos trabalhos utilizaram o ML como procedimento de estimação (61%); a grande maioria dos que não o utilizaram justificaram a decisão em função de características relacionadas à não normalidade dos dados, tanto em relação à curtose, quanto à assimetria.

De forma não surpreendente, nenhum problema de convergência e de soluções inadequadas (valores impossíveis, como variâncias negativas) foi relatado no Brasil, contra apenas um e cinco por cento, respectivamente, no exterior. Talvez isto possa ser explicado pela suposição de que os modelos que não apresentaram convergência ou que resultaram em soluções inadequadas provavelmente não foram submetidos à publicação e, caso tivessem sido, provavelmente encontrariam dificuldades para serem aceitos pelos revisores. A não aceitação seria esperada, não por conta da inadequação teórica dos modelos, mas em função da não convergência decorrente de problemas de especificação dos modelos, das equações ou de problemas diversos relativos à coleta de dados, ao tratamento de *outliers*, aos *missing values* ou, ainda, à entrada de dados.

Em relação à avaliação do ajuste geral do modelo, o índice mais utilizado no Brasil e no exterior tem sido a estatística qui-quadrado ( $\chi^2$ ) de forma absoluta (BR=92%; INT'L=90%). O tamanho médio da razão  $\chi^2$  dividido pelos graus de liberdade (medida de ajuste de parcimônia) foi de 6,81 (BR) e 1,62 (INT'L). Regras práticas indicam que relações até 5 são consideradas satisfatórias (Baumgartner & Homburg, 1996). Neste último caso, os valores brasileiros ficaram substancialmente piores do que os internacionais, um pouco acima, até mesmo do padrão considerado satisfatório. Este é um indicativo de que boa parte dos modelos não atingiu índices de ajustes satisfatórios. Entretanto, como o teste qui-quadrado

é, às vezes, limitado em decorrência de violações da normalidade, e bastante influenciado pelo tamanho da amostra (Bentler, 1990), há a necessidade de se verificarem alguns outros índices, para que se possa ter uma avaliação mais adequada dos modelos testados.

Além da medida de ajuste de parcimônia anteriormente citada ( $\div 2/gl$ ) a **medida de adequação absoluta** mais utilizada é o GFI, com magnitude média pior no Brasil (GFI: BR=0,88; INT'L=0,95). Dentre **medidas de ajuste incrementais**, o índice mais utilizado foi o CFI (69% dos trabalhos no Brasil, contra 13% no exterior), com valores de 0,92 no Brasil e 0,95 no exterior. É interessante notar que o indicador BBI foi utilizado em 23% dos trabalhos no exterior, e em nenhum trabalho no Brasil.

Um destaque negativo nos trabalhos brasileiros pode ser feito em relação às medidas de confiabilidade dos construtos: 25% dos artigos avaliados não apresentaram nenhuma forma de medida de confiabilidade (como Alpha de Cronbach, confiabilidade composta ou outra), contra 22% no exterior.

Uma situação mais aguda foi identificada em relação ao coeficiente de determinação ( $R^2$ ), o qual foi reportado apenas em 33% dos casos no Brasil. Esta situação foi ainda pior no exterior, com apenas 23% de citações dessa medida.

Um dos mais importantes (e também negligenciados) assuntos diz respeito à discussão sobre especificações e reespecificações de modelo. Conforme citam Baungartner e Homburg (1996, p. 155), é pouco provável que um modelo inicialmente especificado seja uma representação plausível dos dados. Ou seja, a reespecificação frequentemente ocorre. Entretanto ela jamais deve ser feita apenas com base nos dados, mas sim através da análise da viabilidade teórica da citada reespecificação. Em outras palavras, qualquer reespecificação deve ser baseada na teoria e não nos dados (Kline, 1998). Obviamente, para que uma reespecificação seja justificada teoricamente, deve haver uma discussão no relato da pesquisa, o que foi feito em apenas 36% dos casos no Brasil, contra 54% no exterior. Por outro lado, a recomendada prática de comparação de modelos rivais é pouco encontrada nos artigos: 36% Brasil, e 31% no exterior. Entretanto, tem-se verificado que, quando há modelos rivais, eles normalmente apresentam sustentação teórica, uma vez que em 31% de todos os artigos analisados no Brasil foram encontradas evidências de fundamentação teórica de modelos rivais.

A partir da análise comparativa entre os trabalhos publicados no ENANPAD e nos principais periódicos internacionais de marketing que utilizaram SEM, pode-se fazer algumas recomendações sobre a sua utilização em trabalhos futuros (vide Tabela 2). Estas recomendações são baseadas em problemas encontrados nas diversas etapas da aplicação de SEM, utilizando-se categorias semelhantes às analisadas por Baungartner e Homburg (1996).

**Tabela 2: Problemas Identificados e Recomendações para os Futuros Trabalhos que Utilizem SEM na Área de Marketing no Brasil**

	<b>Problemas encontrados</b>	<b>Recomendações</b>
Especificação do modelo de mensuração	<p>1) apesar da relação entre o número de variáveis observadas e o número de variáveis latentes ter apresentado um valor aceitável, em diversos trabalhos publicados não foi possível identificar a quantidade de itens por construto.</p> <p>2) alguns trabalhos utilizaram variáveis latentes mensuradas por apenas um item.</p>	<p>1) Incluir no corpo do artigo a matriz de correlação ou de variância/covariância.</p> <p>2) informar com exatidão quantos e quais itens foram utilizados para medir quantos e quais construtos, respectivamente.</p> <p>3) medir cada construto com, pelo menos, três ou quatro itens.</p> <p>4) Detalhar o processo de refino do instrumento: quais critérios foram usados para identificação e descarte de itens dos construtos.</p>
Especificação do modelo estrutural	<p>1) poucos trabalhos relataram correlação entre erros.</p>	<p>1) apesar de os dados não terem sido informados, vários trabalhos podem ter apresentado problemas de correlação entre erros; assim, sugere-se que sempre sejam reportadas correlações entre erros, com as respectivas justificativas teóricas para tal (mas que a correlação seja evitada sempre que possível).</p>
Tamanho de amostra	<p>1) vários estudos não reportaram o número total de parâmetros estimados.</p>	<p>1) apesar de os estudos brasileiros terem apresentado, na média, mais que 10 casos por parâmetro estimado, sugere-se que sejam informados com clareza o tamanho da amostra e, principalmente, o número de parâmetros estimados. Ainda mais importante, que sejam calculados os parâmetros do modelo antes da coleta de dados, para que esse número ajude a determinar o tamanho desejado da amostra.</p>
Identificação do modelo	<p>1) não foram comentados problemas de identificação ou convergência dos modelos, o que pode levar a problemas de estimação de parâmetros, quando ocorrem esses problemas.</p>	<p>1) apresentar os passos seguidos na construção do modelo original, antes de chegar ao teste de modelos rivais (apontando quando foram inseridas correlações, que efeitos geraram nos índices, etc.).</p> <p>2) incluir comentários acerca de eventuais problemas de identificação e convergência de modelos que possam ter surgido, e como foram resolvidos.</p>
Graus de liberdade	<p>1) em vários artigos não foram informadas as contribuições de graus de liberdade oriundas do modelo de mensuração e do estrutural.</p> <p>2) a quantidade de graus de liberdade informada traz alguns indícios de que os modelos tenham sido estimados com construção de parâmetros simultaneamente no item e em seu fator correspondente.</p>	<p>1) informar quanto foi a contribuição do modelo de mensuração e do estrutural para o número total de graus de liberdade.</p> <p>2) informar quais cargas fatoriais foram constritas, e em quais itens ou fatores.</p>

(continuação)  
**Tabela 2: Problemas Identificados e Recomendações para os Futuros Trabalhos que Utilizem SEM na área de Marketing no Brasil**

	<b>Problemas encontrados</b>	<b>Recomendações</b>
Preparação dos dados	1) falta de informação a respeito de preparação dos dados para estimação do modelo, o que prejudica a análise e interpretação dos resultados.	1) reportar como foi a “limpeza” dos dados coletados antes da estimação do modelo, na verificação dos pressupostos necessários para utilização de SEM (ex: tratamento de <i>outliers</i> , <i>missing values</i> e normalidade dos dados).
Mensuração de correlações e associações	1) muitas análises baseadas em correlações.	1) dadas as limitações que as matrizes de correlação podem apresentar, sugere-se a utilização preponderante de matrizes de covariância para entrada de dados.
Estimação do modelo	1) em alguns artigos não foi informado o procedimento de estimação utilizado.	1) sempre informar o procedimento de estimação escolhido, bem como a razão da escolha.
Avaliação do ajuste do modelo	1) Apesar da utilização de índices de ajuste alternativos no Brasil ter sido superior à identificada no exterior, alguns trabalhos ficaram restritos a poucos índices.	1) diversificação de índices de ajuste utilizados, além da relação qui-quadrado/graus de liberdade e GFI.
Avaliação do modelo de mensuração	1) falta de informação a respeito de confiabilidade dos construtos em 25% dos artigos.	1) sempre informar medidas de confiabilidade, como $\alpha$ , confiabilidade composta ou outras. 2) utilizar no mínimo três ou quatro itens para mensuração de construtos.
Avaliação do modelo de estimação	1) falta de informação a respeito da variância medida.	1) sempre informar o coeficiente de determinação ( $R^2$ ).
Análise de resíduos	1) não foram reportadas análises finais do modelo com base em análise residual.	1) incluir análise residual dentre os parâmetros de avaliação do modelo.

Fonte: análises dos dados.

## Considerações Finais, Limitações e Sugestões para Trabalhos Futuros

Um ponto importante a ser destacado, quando se compara a publicação de trabalhos que utilizaram a Modelagem em Equações Estruturais no Brasil e em alguns dos principais periódicos internacionais da área de marketing é a questão temporal de publicação. Enquanto no exterior se pôde identificar trabalhos que utilizam SEM, pelo menos desde a década de 1980 (Baumgartner & Homburg, 1996), no Brasil, considerando os últimos dez anos de produção acadêmica, apenas no final da década de 90 observou-se a utilização mais freqüente de SEM.

Inicialmente, esta diferença temporal nas publicações trouxe algumas expectativas sobre os trabalhos no País. Em primeiro lugar, como os trabalhos brasileiros são mais recentes, seria razoável supor que eles já poderiam ter incorporado uma série de aprimoramentos naturais ocorridos na SEM desde a década de 80, principalmente em relação aos requisitos de especificação de modelos de mensuração e estruturais, preparação de dados, técnicas de estimação de parâmetros e avaliação e interpretação de resultados. Em outras palavras, problemas na utilização de SEM resultantes de novos desenvolvimentos teóricos e metodológicos não só são esperados, mas também bastante prováveis, visto que a construção do conhecimento se dá também através de procedimentos de tentativa e erro.

Em suma, a qualidade da aplicação recente de SEM nos trabalhos de marketing do ENANPAD e das principais publicações acadêmicas nacionais tem sido, em média, semelhante e em alguns critérios até de melhor qualidade do que a aplicação identificada por Baumgartner e Homburg (1996). Entretanto, há alguns pontos específicos que merecem atenção especial dos pesquisadores nacionais. Este artigo destacou alguns destes pontos, e sugeriu algumas rotas de ação. Infelizmente, ainda não houve um esforço semelhante para mapeamento do uso recente de SEM no cenário internacional. Fica, portanto, esta idéia como importante sugestão para pesquisas futuras.

**Artigo recebido em 20.07.2004. Aprovado em 05.11.2004.**

## REFERÊNCIAS BIBLIOGRÁFICAS

Anderson, J. C.,  
Gerbing, D. W. (1988, May).  
Structural equation modeling in

practice: A review and recommended  
two-step approach. *Psychological  
Bulletin*, 103(3), 99-111.

- Bagozzi, R. P. (1980). *Causal models in marketing*. New York: Wiley.
- Bagozzi, R. P., & Baumgartner, H. (1994). The evaluation of structural equation models and hypothesis testing. In R. Bagozzi *Principles of marketing research*. Cambridge: Blackwell.
- Baumgartner, H., & Homburg, C. (1996, April). Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing*, 13(2), 139-161.
- Bentler, P. (1980). Multivariate analysis with latent variables: causal modeling. *Annual Review of Psychology*, 31, 419-456.
- Bentler, P., & Chou, C. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research*, 16, 78-117.
- Bentler, P. (1990, March). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246.
- Bollen, K. (1989). *Structural equation modeling with latent variables*. New York: Wiley.
- Byrne, B. (1994). *Structural equation modeling with EQS and EQS/Windows: basic concepts, applications and programming*. Thousand Oaks: Sage.
- DeVellis, R. F. (1991). *Scale Development (Applied Social Research Methods Series)*. Newbury Park: Sage Publications.
- Gujarati, D. (2000). *Econometria básica*. São Paulo: Makron.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River: Prentice-Hall.
- Hulland, J., Chow, Y., & Lam, S. (1996, April). Use of causal models in marketing research: a review. *International Journal of Research in Marketing*, 13(2), 181-197.
- Jarvis, C. B., Mackenzie, S., & Podsakoff, P. (2003, September). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218.
- Joreskog, K.G., & Sorbom, D. (1993). *LISREL 8: structural equation modeling with the simplis command language*. Chicago: Scientific Software International.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

- L. R. James, &  
L.A. James (1989).  
Causal modeling in organizational research. In C. Cooper, I. Robertson (Eds.). *International Review of Industrial and Organizational Psychology*. New York: Wiley.
- Maccallum, R., &  
Austin, J. (2000).  
Applications of structural equation modeling. *Psychological Research*, 51(1), 201-226.
- Medsker, G,  
Williams, L., &  
Holohan, L. (1994, Summer).  
A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20(2), 439-464.
- Pedhazur, E., &  
Schmelkin, L. (1991).  
*Measurement, design, and analysis: an integrated approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Steenkamp, J., &  
Trijp, H. (1991, November).  
The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing*, 8(4), 283-299.