

Abrindo a caixa preta: considerações sobre a utilização da Análise Fatorial Confirmatória nas pesquisas em Administração

Opening the black box: considerations on the use of Confirmatory Factor Analysis in research in Business Management

Humberto Elias Garcia Lopes

RESUMO

A Análise Fatorial Confirmatória (AFC) é uma técnica multivariada que vem sendo adotada com relativa frequência nas pesquisas em Administração. Entretanto, para que ela seja utilizada corretamente, é preciso que se tenha domínio de seus aspectos conceituais e operacionais. Este artigo tem por objetivo oferecer aos pesquisadores da área um roteiro para a condução de uma investigação que utilize a AFC. Por meio da revisão da literatura sobre o assunto e da utilização de um exemplo empírico, este artigo procurou também mostrar os principais problemas que ocorrem nesse tipo de análise. Embora as dificuldades inerentes à técnica não inviabilizem o uso de tal análise, é preciso reconhecer que elas têm que ser consideradas como limitações significativas em uma pesquisa, tirando da técnica a possibilidade de ser encarada como uma solução universal para qualquer problema científico. Dessa forma, é aconselhável que o pesquisador planeje criteriosamente a sua investigação, ponderando os prós e os contras da estratégia de análise proposta, a fim de que a escolha feita seja realmente adequada à pesquisa.

Palavras-chave: Análise Fatorial Confirmatória; Valores discrepantes; Dados ausentes; Análise multivariada; Pressupostos.

Em anos recentes, tem sido freqüente a utilização de sofisticadas metodologias e técnicas estatísticas multivariadas nas pesquisas em Administração, fato que não é exatamente novo, já que sua origem data dos anos 1970 e 1980 (AAKER, 1971; BISHOP, 1995; CHATFIELD; COLLINS, 1980; MUIRHEAD, 1982; JOHNSON; WICHERN, 2002). Entre as diversas técnicas disponíveis, uma tem-se destacado em estudos nacionais recentes na área: a Análise Fatorial Confirmatória (AFC).

• Artigo recebido em 13/07/2004 e aprovado para publicação em 15/10/2005.

Atualmente, há significativo número de pesquisas no País que utilizam a AFC, como mostram os trabalhos de Urdan e Rodrigues (1999), Latif (2000), Veiga (2000), Urdan e Zuñiga (2001), Filho (2001), Lopes (2001, 2003) e Gosling (2004). No entanto, embora seja uma técnica multivariada que permite uma análise complexa em pesquisas quantitativas, ela tem particularidades que podem se transformar em problemas consideráveis para o pesquisador, comprometendo a qualidade da investigação. Para que esse risco seja reduzido a níveis aceitáveis, é preciso dominar a técnica, tanto em seus aspectos teóricos como operacionais.

O objetivo deste trabalho é fornecer aos pesquisadores uma proposta de roteiro para a utilização da AFC nas investigações em Administração. Para que esse esforço resultasse em algo consistente e útil, a abordagem aqui utilizada procurou dar especial ênfase às limitações da técnica, cuja observância é fundamental para o sucesso da investigação.

ASPECTOS CONCEITUAIS

O primeiro aspecto conceitual a ser salientado é que a Análise Fatorial Confirmatória é um caso especial de uma técnica mais ampla, denominada modelagem de equações estruturais (HAIR JR. et al. 1998; ULLMAN, 2001; MALHOTRA, 2001). Nela, busca-se descrever os relacionamentos existentes entre dois tipos de variável: as "latentes", definidas como "um conceito hipotetizado e não observado, do qual se pode aproximar apenas através de variáveis observáveis ou mensuráveis" (HAIR JR. et al., 1998, p. 585), e as "manifestas", também denominadas "indicadores", que consistem em valores observados e que são usados como medida de um conceito ou construto (HAIR JR. et al., 1998, p. 580). No caso das variáveis latentes, é preciso destacar ainda que elas podem ser "exógenas", quando funcionam apenas como preditoras de outros construtos, ou "endógenas", que são as variáveis dependentes em pelo menos um relacionamento causal.

Mueller (1996) destaca ainda que a Análise Fatorial Confirmatória está baseada na premissa de que as variáveis observadas são indicadores imperfeitos de certos construtos latentes. Dessa forma, se mais de um indicador é usado para medir um construto específico, a AFC permite ao pesquisador agrupar tais indicadores em maneiras pré-especificadas, a fim de avaliar em que extensão determinado conjunto de dados aparentemente confirma a estrutura prevista.

Os relacionamentos entre esses dois tipos de variável são especificados por meio de diagramas de caminho, nos quais os construtos latentes são representados por elipses, e os indicadores, por retângulos. Associados a esses estão setas indicadoras da presença de erros de mensuração. Isso ocorre porque o pesquisa-

dor deve admitir que os dados foram coletados e tratados de maneira imperfeita. A Figura 1 mostra um exemplo de diagrama de caminho.

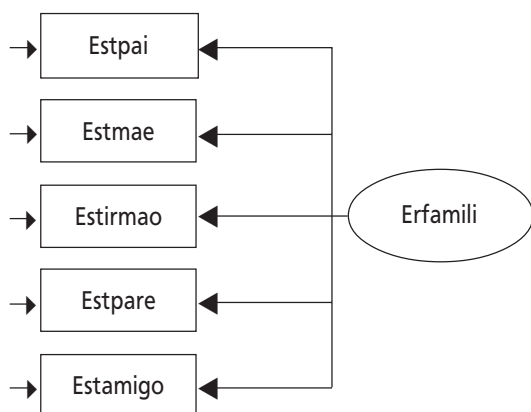


Figura 1: Exemplo de diagrama de caminho.

Fonte: Adaptado de Lopes (2001).

No diagrama exibido pela Figura 1, há o construto "Erfamili", que é mensurado por meio de cinco indicadores: "Estpai", "Estmae", "Estirmao", "Estpare" e "Estamigo". As setas à esquerda dos indicadores representam os erros de mensuração associados ao modelo.

A definição do relacionamento entre as variáveis é crucial para a correta utilização da AFC, posto que explicita a função operacional de cada uma delas, ou seja, qual será a variável independente e qual será a dependente. Sobre esse ponto, é preciso destacar que uma das vantagens da AFC está justamente na possibilidade de se ter diversas variáveis dependentes e independentes, o que permite testar múltiplos relacionamentos ao mesmo tempo.

ASPECTOS OPERACIONAIS DA ANÁLISE FATORIAL CONFIRMATÓRIA

A condução de uma investigação baseada em uma técnica de análise multivariada com alto nível de complexidade exige cuidados do pesquisador. Antes mesmo de utilizar a AFC propriamente dita, há diversas etapas que devem ser cumpridas, sem as quais os resultados poderão conter vieses significativos, enfraquecendo os achados da pesquisa.

Uma consequência disso é que o pesquisador, ao utilizar a AFC ou qualquer outra técnica de análise multivariada, deve ter a consciência de que a pesquisa talvez não seja conduzida com a rapidez desejada, apesar dos sofisticados recursos computacionais disponíveis atualmente. Na verdade, *softwares* e computadores

apenas processam as informações que lhes são repassadas. Portanto, se essas apresentam vícios em sua origem, os resultados não serão os melhores. Justamente para auxiliar o pesquisador a evitar tal problema, a seguir são descritas as etapas na condução da Análise Fatorial Confirmatória.

A escolha do instrumento de coleta dos dados

A definição do tipo de instrumento que será utilizado na coleta de dados é importante para uma pesquisa consistente. Em princípio, tanto o questionário como a entrevista, instrumentos tradicionais de coleta de dados, podem ser adequados a uma investigação científica. No caso da AFC, porém, o pesquisador deve considerar a necessidade de coletar informações que expressem relações quantitativas. Esse pode ser um aspecto a favor do questionário, dada a sua capacidade de registrar uma massa significativa de dados de um grande número de respondentes (SELLTIZ; WRIGHTSMAN; COOK, 1987). Essa característica é importante para a AFC, uma vez que ela demanda quantidade expressiva de dados, a fim de que o modelo fatorial chegue a uma solução única. Na verdade, há uma recomendação para que se tenham entre 10 e 15 respondentes para cada indicador incluído na pesquisa (HAIR JR. *et al.*, 1998).

A decisão sobre qual o instrumento mais apropriado pode ser facilitada por meio do trabalho de Silva e Menezes (2000). Esses autores sugerem um sistema de classificação de pesquisas com base em quatro critérios: natureza, forma de abordagem do problema, objetivos e procedimentos técnicos. Tal sistema pode ser útil para o pesquisador melhor delinear a sua investigação, o que pode ajudá-lo a escolher o instrumento mais adequado para a coleta dos dados.

A definição das escalas

A especificação do tipo de escala utilizada nos indicadores da pesquisa é fundamental para a AFC. Em termos teóricos, essa técnica é capaz de lidar tanto com escalas contínuas como discretas (HAIR JR. *et al.*, 1998). Entretanto, na prática, incorporar escalas nominais ou ordinais - discretas por definição - implica dificuldades adicionais a ser superadas pelo pesquisador, já que na AFC os dados coletados mediante tais escalas são incluídos nos cálculos.

Apesar disso, é preciso reconhecer que dificilmente um pesquisador em Administração conseguirá conduzir uma investigação de natureza quantitativa sem utilizar escalas discretas, especialmente as ordinais. Assim, torna-se necessário estabelecer estratégias que permitirão a melhor utilização desse tipo de escala na AFC.

Sobre isso, uma primeira recomendação ao pesquisador é utilizar as escalas

"nominais" somente para descrição dos dados. Conforme mostram Nunnally e Bernstein (1994), por nem sempre refletirem relacionamentos quantitativos, essas escalas têm propriedades matemáticas limitadas. Assim, prestam-se melhor à descrição do perfil ou à definição de agrupamentos da amostra utilizada, por exemplo. Por outro lado, as escalas "ordinais" estão em um patamar mais alto de mensuração, já que conseguem medir grandezas que podem ser traduzidas em termos quantitativos. Mesmo assim, suas propriedades matemáticas também apresentam limitações consideráveis, que vão exigir atenção do pesquisador. Dados medidos por meio de escalas ordinais podem ser analisados por estatísticas não-paramétricas, tais como medianas, quartis e a correlação de Spearman.

Uma terceira opção é a escala "intervalar", na qual são refletidas operações que definem uma unidade de mensuração em termos de relacionamento de igualdade ou de desigualdade, conforme observado por Nunnally e Bernstein (1994) e Malhotra (2001). Definida como contínua, a escala intervalar tem um zero arbitrário e permite a utilização de estatísticas paramétricas, como a média aritmética, a variância e a correlação de Pearson. Ainda que haja limites em suas propriedades matemáticas, a escala intervalar pode ser utilizada na AFC com menos restrições que aquelas inerentes às escalas discretas.

O quarto tipo de escala é denominado "razão", que consiste em uma escala intervalar com um zero racional e não arbitrário. Ela permite um número considerável de operações matemáticas, possibilitando o uso de medidas paramétricas. Dessa forma, a escala razão é bastante útil em pesquisas quantitativas, apesar de nem sempre ser possível ou viável adotá-la na Administração.

Embora as escalas intervalares ou razão sejam, por vezes, difíceis de ser utilizadas nas pesquisas em Administração, há um artifício que permite que uma escala ordinal possa ser tratada como intervalar, o que auxilia significativamente a AFC. Essa estratégia consiste em utilizar determinado número de categorias na escala, de tal forma que a perda de informação ao tratá-la como intervalar seja reduzida, ainda que suas propriedades matemáticas continuem sendo típicas de uma escala ordinal.

Sobre esse aspecto, é preciso destacar que não existe consenso sobre qual o número exato de categorias que permitiria a uma escala ordinal ser tratada como intervalar. Por exemplo, Jöreskog e Sörbom (2000b) estabelecem que isso só seria possível em uma escala com quinze ou mais categorias. Já Nunnally e Bernstein (1994) afirmam, com base em pesquisas empíricas, que uma escala com onze categorias não provoca perda significativa de informações. Na verdade, para esses autores, até mesmo escalas entre sete e nove categorias poderiam ser tratadas como contínuas. Qualquer que seja a opção do pesquisador, é preciso considerar que haverá alteração na precisão e na acurácia da escala, conforme explica Pereira (1999).

Para a utilização da AFC, recomenda-se que o pesquisador inclua no diagrama de caminho somente indicadores que possam ser medidos por meio de escalas contínuas ou ordinais, desde que essas possam ser tratadas como intervalares. Do contrário, os resultados da análise poderão conter deficiências que comprometerão toda a pesquisa.

A análise preliminar dos dados

Antes que qualquer técnica mais complexa de análise quantitativa possa ser utilizada, é necessário que o pesquisador conheça os dados que coletou (HAIR JR. *et al.*, 1998; TABACHNICK; FIDELL, 2001). Com isso, torna-se possível descrever o perfil da amostra estudada, bem como detectar eventuais falhas de tabulação dos dados, o que reduz o risco de erros de análise.

Para que essa análise seja conduzida satisfatoriamente, além do uso de estatísticas para a descrição do perfil da amostra e das frequências obtidas, o pesquisador deverá recorrer a duas técnicas. A primeira é o estudo dos dados ausentes (*missing data*), e a segunda, a identificação dos valores discrepantes (*outliers*), tanto univariados como multivariados.

Em relação à primeira técnica, é preciso considerar que os dados ausentes podem ser definidos como uma informação não disponível a respeito de uma pessoa ou caso sobre os quais outras informações estão disponíveis (HAIR JR. *et al.*, 1998). Esse tipo de dado surge frequentemente nas pesquisas, uma vez que as dificuldades inerentes à coleta de dados tornam pouco provável que o pesquisador não tenha que desenvolver estratégias para lidar com os *missing data* (HAIR JR. *et al.*, 1998; TABACHNICK; FIDELL, 2001; MALHOTRA, 2001).

Uma dificuldade gerada pelos dados ausentes é que eles podem trazer vieses nos resultados da pesquisa, se forem excluídos ou substituídos sem análise prévia da sua aleatoriedade. Assim, o pesquisador corre o risco de retirar ou substituir informações que apresentam um padrão de distribuição na matriz de dados utilizada, comprometendo as análises subsequentes.

Mesmo que sejam aleatórios, os dados ausentes devem ser tratados criteriosamente, já que existem dois níveis diferentes de aleatoriedade. Hair Jr. *et al.* (1998) afirmam que os dados podem ser aleatórios ao acaso (AAA) ou completamente aleatórios ao acaso (CAAA), sendo este o maior grau de aleatoriedade ao qual é possível chegar. Segundo esses autores, para detectar se os dados são AAA, é necessário fazer o teste "t" de comparação das médias de dois grupos: um formado pelos casos sem dados ausentes, e o outro, pelos casos com dados ausentes. Esse teste é repetido para cada indicador do estudo.

Além do teste "t", é preciso utilizar correlações bivariadas entre indicadores

dicotomizados. Esses são formados ao se substituir os valores válidos por 1, e os dados ausentes, por 0. As correlações mostram o grau no qual os dados ausentes estão relacionados aos pares de indicadores. Baixas correlações indicam baixa associação entre o processo que levou à existência de dados ausentes e os dois indicadores considerados, e vice-versa.

Por outro lado, a verificação dos dados CAAA é realizada por meio de um teste global que faz a comparação entre o padrão real de dados ausentes e aquele que seria esperado se tais dados fossem distribuídos totalmente ao acaso. Uma vez que o pesquisador somente poderá ter a garantia de que a exclusão dos dados ausentes não prejudicará a análise de seus dados se eles forem CAAA, o teste global é importante para garantir maior rigor no estudo.

No que se refere à segunda técnica, define-se um valor discrepante como uma observação que é essencialmente distinta de outras observações (HAIR JR. *et al.*, 1998). De acordo com Tabachnick e Fidell (2001), ele precisa ser identificado na pesquisa porque sua presença pode levar a uma distorção significativa dos resultados. Os valores discrepantes podem ser detectados por meio de análises univariadas, bivariadas ou multivariadas. Na primeira, o que se busca é identificar um valor consideravelmente extremo em uma variável. Uma das maneiras para que isso seja feito é utilizar o diagrama em caixas (*boxplot*), no qual os valores discrepantes são aqueles situados a 1,5 ou mais desvios quartílicos do quartil superior ou inferior da distribuição.

Outra maneira é o cálculo dos escores *z* das variáveis, no qual os valores originais são convertidos em escores padronizados, que têm média 0 e desvio-padrão 1. Os escores são então comparados com um valor-limite, obtido em função do tamanho da amostra utilizada e do nível de significância desejado. No caso de amostras grandes, que contenham mais de 80 observações, Hair Jr. *et al.* (1998) sugerem um valor entre 3 e 4. Entretanto, Tabachnick e Fidell (2001) são mais específicas, propondo que, com significância de 0,001 em um teste bilateral, escores "*z*" fora do intervalo [-3,29; 3,29] sejam considerados como valores discrepantes. Por levar em conta um número maior de parâmetros do que os diagramas em caixa, o que pode resultar em maior precisão e rigor, a análise dos escores "*z*" pode ser mais adequada para detectar os valores discrepantes univariados em uma pesquisa.

A análise bivariada dos valores discrepantes é realizada através de diagramas de dispersão, nos quais se analisa a distribuição dos dados entre pares de indicadores contínuos (HAIR JR. *et al.*, 1998). Todavia, por depender significativamente da avaliação subjetiva do pesquisador, essa opção pode levar a menor precisão dos resultados.

Tabachnick e Fidell (2001) consideram que os valores discrepantes multivariados são resultantes de uma combinação anormal de escores em dois ou mais indi-

cadores. Dessa forma, para detectá-los, uma das técnicas disponíveis é a distância de Mahalanobis (D2), definida por essas autoras como a distância de um caso da centróide dos casos remanescentes, em que a centróide é o ponto criado pela interseção das médias de todas as variáveis consideradas. Para utilizar a D2, é preciso definir o nível de significância desejado e o número de graus de liberdade utilizados, a fim de que se tenha um valor-limite, fornecido pela distribuição qui-quadrado (χ^2), acima do qual o caso deve ser considerado como um valor discrepante multivariado. Contudo, Mueller (1996) salienta que os testes baseados na distribuição qui-quadrado geralmente são rejeitados quando o tamanho da amostra é superior a 200, levando, em alguns casos, à necessidade de se fazer testes adicionais. Em relação ao primeiro parâmetro, tanto Hair Jr. *et al.* (1998) como Tabachnick e Fidell (2001) recomendam um nível conservador, de 0,001. Por sua vez, o segundo parâmetro é dado pelo número de indicadores contínuos considerados.

Após identificados os valores discrepantes multivariados, é preciso descobrir os indicadores que provocaram esse resultado, o que é feito por meio da análise de regressão stepwise, na qual cada caso com valor discrepante multivariado é avaliado separadamente. Para isso, é criada uma variável dicotômica, cujo valor 0 é atribuído ao caso que contém o valor discrepante, e 1, aos demais. Com o primeiro fixado como dependente e os demais como independentes, a regressão é efetuada.

Uma vez identificados os valores discrepantes multivariados e as variáveis que os provocam, o pesquisador deve decidir pela manutenção ou exclusão dos casos. Se eles forem realmente aberrantes e não representativos das observações na população considerada, devem ser excluídos. Do contrário, pode ser mais adequado mantê-los.

A verificação dos pressupostos da análise multivariada

Sendo a AFC uma técnica de análise multivariada, torna-se necessário verificar o atendimento aos pressupostos da normalidade, linearidade e homogeneidade de variâncias, além de detectar se há problemas em relação à multicolinearidade e à singularidade (TABACHNICK; FIDELL, 2001; MALHOTRA, 2001). Caso isso não seja feito, o pesquisador corre o risco de basear suas conclusões em resultados equivocados.

De todos os pressupostos da AFC, o da normalidade é o mais importante (HAIR JR. *et al.*, 1998). Ele se refere ao formato da distribuição de determinada variável contínua, utilizando como parâmetros o grau de assimetria e curtose. A normalidade univariada pode ser verificada por um exame visual da distribuição dos dados em uma variável, através dos gráficos Q-Q ou P-P. Outra opção é o teste de

Kolmogorov-Smirnov. Ele verifica a hipótese de que os dados considerados constituem uma amostra aleatória de uma distribuição específica, entre elas, a normal. Contudo, Kanji (1999) ressalta que esse teste tem como limitação o fato de ser aplicável somente quando a função de distribuição da população for contínua. Uma terceira opção é calcular a estatística z para a assimetria e curtose da distribuição. Se qualquer um desses valores se situar fora do intervalo $[-1,96; 1,96]$, a distribuição é não-normal, supondo um nível de significância de 0,05.

Violações à normalidade univariada não impedem a análise, mas podem enfraquecer as associações entre os indicadores. Exatamente por isso existem métodos disponíveis para se tentar minimizar a assimetria e a curtose de variáveis não-normais. No entanto, mesmo assim, nem sempre uma variável consegue atingir a normalidade. Diversas transformações podem ser necessárias para que a distribuição se torne menos assimétrica e com grau menor de curtose. Segundo Graf (1999), isso pode se tornar um problema significativo para a pesquisa, uma vez que as transformações, na verdade, geram variáveis com valores diferentes daqueles observados empiricamente, o que implica uma limitação.

Contudo, Jöreskog *et al.* (2000) sugerem uma estratégia que pode reduzir tais problemas. Ela consiste na normalização dos indicadores, transformando-os em escores "z". Com isso, suas origens e unidades de mensuração passam a não ter significado intrínseco, o que torna as estatísticas menos sensíveis às violações à normalidade. Caso a normalização reduza a assimetria e a curtose, os dados que sofreram essa transformação passam a ser utilizados para os demais testes de verificação dos pressupostos da normalidade. A Tabela 1 e a Tabela 2 mostram um exemplo das alterações proporcionadas por essa estratégia em um construto denominado de "estrutura de relacionamentos familiares e de amizade".

Tabela 1. Estrutura de relacionamentos familiares e de amizade – teste de normalidade univariada com dados não normalizados.

Indicador	Assimetria		Curtose		Assimetria e curtose	
	Escore z	Valor p	Escore z	Valor p	Qui-quadrado	Valor " p "
Pai	-4,40	<0,00	-0,29	0,78	19,45	<0,00
Mãe	-6,41	<0,00	0,96	0,34	41,99	<0,00
Irmão ou irmã	-6,27	<0,00	0,84	0,40	39,96	<0,00
Outros parentes	-1,60	0,11	-6,51	<0,00	44,97	<0,00
Amigos	-6,23	<0,00	1,39	0,17	40,69	<0,00

Fonte: Adaptado de Lopes (2001).

A Tabela 1 contém o resultado do teste de normalidade univariada com os dados originais. Ela comprova que nenhum dos indicadores atendeu ao pressuposto da normalidade univariada, já que o valor " p " para a estatística qui-quadrado

de todos eles foi inferior ao nível de 0,05, adotado no teste. No entanto, quando se opta por trabalhar com escores já normalizados, os resultados são diferentes, conforme mostra a Tabela 2.

Tabela 2. Estrutura de relacionamentos familiares e de amizade – teste de normalidade univariada com dados normalizados.

Indicador	Assimetria		Curtose		Assimetria e curtose	
	Escore z	Valor p	Escore z	Valor p	Qui-quadrado	Valor "p"
Pai	-1,45	0,15	-0,95	0,34	3,01	0,22
Mãe	-1,77	0,08	-0,53	0,60	3,40	0,18
Irmão ou irmã	-1,27	0,20	-0,94	0,35	2,50	0,29
Outros parentes	0,02	0,98	-1,46	0,14	2,14	0,34
Amigos	-1,00	0,32	-0,63	0,53	1,40	0,50

Fonte: Adaptado de Lopes (2001).

Nota: O valor "p" de assimetria e curtose testa a hipótese nula de que a distribuição do indicador é normal. Foi utilizada uma significância de 0,05 em um teste bilateral.

Como evidencia a Tabela 2, a normalização fez com que todos os indicadores passassem a atender ao pressuposto da normalidade univariada. Isso significa que houve melhora significativa na distribuição, o que pode se traduzir em benefícios para a AFC.

Quanto à normalidade multivariada, é possível verificá-la através de um teste, no qual são calculados os escores "z" para assimetria e curtose do conjunto de variáveis contínuas consideradas. Distribuições multivariadas nas quais esses dois parâmetros apresentam alto e significativo grau dificilmente atentam ao pressuposto da normalidade. Na verdade, violações à normalidade univariada já tornam pouco provável que as variáveis atendam à normalidade multivariada (VEIGA, 2000). Para que haja essa confirmação, o teste também calcula a estatística qui-quadrado para assimetria e curtose multivariada das variáveis. A hipótese nula é a de que a função distribuição das variáveis contínuas consideradas provém de uma distribuição normal em termos multivariados, com nível de significância de 5% (JÖRESKOG *et al.*, 2000). Assim, valores "p" inferiores a 0,05 rejeitam essa hipótese.

Por sua vez, segundo Hair Jr. *et al.* (1998), a linearidade é um pressuposto de todas as técnicas multivariadas de análise baseadas em medidas de associação. Dessa forma, é importante verificar se as variáveis consideradas apresentam um grau significativo de linearidade. Isso pode ser feito através de diagramas de dispersão, nos quais são tomados os relacionamentos bivariados entre as variáveis.

Embora seja útil, essa técnica demanda expressiva subjetividade do pesquisador na avaliação do grau de linearidade, o que pode se tornar uma limitação da pesquisa. Uma alternativa é a utilização das correlações bivariadas, que produzem resultados menos dependentes de análises subjetivas. Nesse caso, os parâ-

metros são os coeficientes de correlação de Pearson ou Spearman. Esse é uma versão não paramétrica do "r" de Pearson, sendo adequado quando há variáveis medidas por meio de escalas ordinais ou quando as variáveis intervalares não atenderam ao pressuposto da normalidade (SPSS, 1999b). Esse coeficiente situa-se no intervalo $[-1; 1]$ e sua interpretação é feita da mesma maneira que o "r" de Pearson. Cabe ao pesquisador determinar o valor mínimo do coeficiente para que seja atendido o pressuposto da linearidade. Sugere-se adotar como ponto de corte valores iguais ou superiores a 0,70 (LOPES, 2003).

A homogeneidade de variâncias se refere ao pressuposto de que as variáveis dependentes apresentam níveis equivalentes de variância ao longo do espectro de variáveis predictoras. Normalmente, esse pressuposto é testado somente quando o estudo compara dois ou mais grupos amostrais. Caso contrário, ou seja, quando se lida com dados não agrupados, a homogeneidade de variâncias admite que a variabilidade nos escores de uma variável contínua é exatamente a mesma para todos os valores de outra variável contínua.

Em relação à multicolinearidade e à singularidade, Tabachnick e Fidell (2001) afirmam que elas consistem em problemas que ocorrem quando as variáveis estão excessivamente correlacionadas. Havendo multicolinearidade, há alta correlação. Ocorrendo singularidade, as variáveis são redundantes, visto que uma se constitui na combinação de duas ou mais das outras. Para verificar se existe singularidade, essas autoras sugerem que a técnica de análise multivariada utilizada seja rodada no computador. Segundo elas, com exceção das técnicas centradas na análise de componentes principais, todas as outras não chegam a nenhum resultado se houver singularidade, fazendo com que o computador emita mensagens de erro.

Já a multicolinearidade pode ser avaliada mediante o recurso *collinearity diagnostics*, oferecido por pacotes estatísticos. O programa agrupa os dados em "dimensões", definidas com base no seu autovalor. Para cada uma delas, é distribuída determinada quantidade de variância entre os indicadores considerados. Se uma dimensão detiver dois ou mais indicadores com variância superior a 0,50 e se seu índice de condição for muito próximo ou superior a 30, há forte indício de presença de multicolinearidade. Nesse caso, é recomendável a exclusão da variável. A Tabela 3 apresenta um exemplo desse procedimento.

A Tabela 3 evidencia que nenhuma das dimensões apresentaram índice de condição significativamente próximo a 30. Além disso, não foram detectadas duas ou mais proporções de variância iguais ou superiores a 0,50 em uma mesma dimensão. Portanto, é possível concluir que, no exemplo dado, a multicolinearidade se restringiu a padrões aceitáveis para a análise multivariada.

Tabela 3. Estrutura de relacionamentos familiares e de amizade – teste de multicolinearidade.

Dimensão	Autovalor	Índice de condição	Proporções de variância					
			Constante	Pai	Mãe	Irmão ou irmã	Outros parentes	Amigos
1	0,462	3,126	0,00	0,00	0,00	0,00	0,25	0,41
2	0,395	3,380	0,00	0,03	0,04	0,00	0,38	0,13
3	0,362	3,531	0,00	0,00	0,03	0,00	0,24	0,08
4	0,212	4,616	0,00	0,00	0,40	0,74	0,06	0,00
5	0,122	6,077	0,00	0,93	0,48	0,21	0,00	0,00

Fonte: Adaptado de Lopes (2001).

Nota: As variâncias superiores a 0,50 estão em negrito.

A Análise Fatorial Confirmatória

Uma análise dos procedimentos adotados por Hair Jr. *et al.* (1998), Jöreskog *et al.* (2000), Tabachnick e Fidell (2001) e Malhotra (2001) leva à conclusão que, em síntese, é possível conduzir satisfatoriamente a AFC via o cumprimento de nove etapas. A primeira é a escolha da matriz de dados a ser utilizada, na qual existem duas opções: a matriz de correlações e a matriz de covariâncias, e cada uma tem vantagens e limitações que precisam ser consideradas pelo pesquisador.

A segunda etapa da AFC é a escolha do método de estimação. Todavia, a não ser que haja alguma particularidade que recomende o contrário, sugere-se que o pesquisador recorra à máxima verossimilhança (*maximum likelihood*), uma vez que ela permite estimativas mais precisas, por ser sensível às violações da normalidade, como salientam Ullman (2001), Schumacker e Beyerlein (2000), Olsson *et al.* (2000) e Jöreskog *et al.* (2000).

A terceira etapa da análise é verificar a identificação do modelo. Em termos gerais, um modelo é identificado quando ele é capaz de gerar estimativas únicas, o que está baseado no princípio de que é preciso ter uma equação única e separada para estimar cada coeficiente.

A quarta etapa é encontrar eventuais estimativas discrepantes (*offending estimates*) nos resultados da AFC. Segundo Hair Jr. *et al.* (1998), as estimativas mais comuns são: variâncias-erro negativas ou não significativas, estimativas padronizadas que excedem o valor de 1,0 e erros padronizados muito grandes, associados a qualquer coeficiente estimado.

A identificação de estimativas não significativas é feita pela exibição dos valores do teste "t" no diagrama de caminho. Conforme sugere Ullman (2001), as propriedades estatísticas da máxima verossimilhança tornam recomendável que o teste seja feito para níveis de significância mais conservadores, de 0,025 ou 0,01, o que demanda valores de "t" situados fora do intervalo [-2,56; 2,56], em um teste bilateral.

A quinta etapa é a avaliação do ajuste global do modelo. Essa é feita pelas medidas de ajuste, calculadas por pacotes estatísticos como o LISREL, o EQS ou o AMOS. É preciso ressaltar que a escolha das medidas utilizadas pode ser complexa, dado o grande número de opções disponíveis. Todavia, três medidas são comumente empregadas, conforme mostra o Quadro 1.

Quadro 1. Medidas freqüentemente utilizadas na Análise Fatorial Confirmatória.

Medida	Características e limitações	Valores recomendados
Qui-quadrado absoluto	Vem acompanhado do valor "p". A hipótese nula é a de que não existe diferença entre a matriz de dados e a matriz estimada. Para amostras acima de 200 elementos, torna-se muito sensível às violações da normalidade, podendo fornecer estimativas distorcidas.	Não há. O ideal é que a hipótese nula seja aceita para um nível de significância.
Raiz quadrada da média do erro de aproximação (RMSEA)	Mede a discrepância dos dados, ajustando-a aos graus de liberdade.	Valores abaixo de 0,08.
Qui-quadrado normalizado	Ajusta o qui-quadrado absoluto aos graus de liberdade.	Valores entre 1,0 e 3,0. Se utilizado um critério mais liberal, o limite superior pode ser de 5,0.

Fonte: Adaptado de Hair Jr. *et al.* (1998), Ulmann (2001) e Mueller (1996).

Nota: O valor "p" de assimetria e curtose testa a hipótese nula de que a distribuição do indicador é normal. Foi utilizada uma significância de 0,05, em um teste bilateral.

Destaca-se, entretanto, que os valores recomendados para cada medida podem variar, dependendo do autor escolhido como referência. Na realidade, o importante é que o pesquisador escolha as medidas e os valores que sejam mais adequados às características da pesquisa que estiver conduzindo.

A sexta etapa é a quantificação do ajuste do modelo de mensuração. Para isso, são utilizados os valores dos coeficientes de determinação (R²) entre o indicador e a variável latente à qual ele se conecta. Quanto maior o valor de R², melhor é o ajuste, visto que isso indica que uma porção mais expressiva das variações no indicador pode ser atribuída a alterações na variável latente.

A sétima etapa é a avaliação do ajuste do modelo estrutural, na qual novamente pode ser utilizado o coeficiente de determinação, com a diferença que ele é extraído das equações que expressam a relação entre as variáveis latentes. Dessa forma, valores mais altos são desejados, pelas mesmas razões descritas na sexta etapa.

A oitava etapa consiste na avaliação da matriz de resíduos normalizados. Hair Jr. *et al.* (1998) recomendam que no máximo 5% dos resíduos estejam fora do intervalo [-2,58; 2,58], o que implica nível de significância de 0,05. Violações a essa recomendação indicam deficiências no ajuste do modelo estimado.

A nona etapa é a análise das estimativas padronizadas que constam do diagrama de caminho. Valores mais próximos dos limites superior e inferior do intervalo $[-1; 1]$ indicam a presença de relacionamentos mais fortes. Como é possível notar, essa interpretação é semelhante à adotada no coeficiente de correlação de Pearson.

UM EXEMPLO DE APLICAÇÃO DA AFC EM UMA PESQUISA EM ADMINISTRAÇÃO

Para este trabalho foi feito um estudo no qual se procurou detectar em que grau um conjunto de variáveis latentes afetava o montante de recursos sociais acumulados por profissionais de nível superior mediante sua rede de contatos pessoais. Entretanto, para que essa proposta fosse viabilizada por meio da AFC, testou-se cada uma das variáveis latentes, a fim de verificar a adequação desses construtos à pesquisa.

Por uma questão de objetividade, foi escolhida a variável latente "estrutura de relacionamentos familiares e de amizade", que identificava qual era a composição da rede de relacionamentos dos profissionais pesquisados. Os indicadores propostos para medir essa variável foram analisados de acordo com os procedimentos explicitados neste trabalho, a fim de que somente aqueles que se mostrassem relevantes fossem associados ao referido construto.

O estudo dos dados ausentes evidenciou que os indicadores propostos apresentavam padrão AAA, ou seja, as informações ausentes eram aleatórias ao acaso. O teste global também apontou para o mesmo resultado, o que autorizou a exclusão dos dados ausentes.

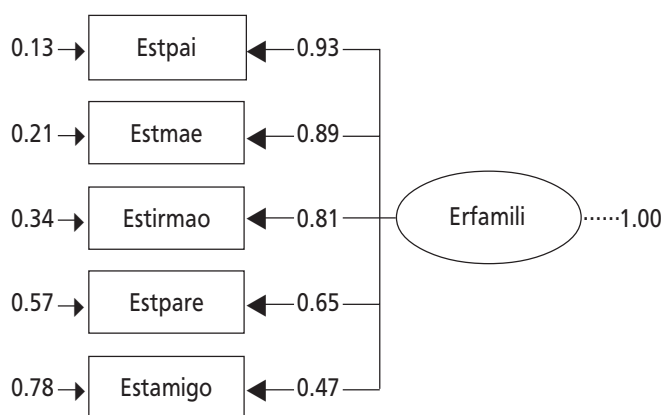
Os valores discrepantes univariados e multivariados foram identificados através dos escores "z" e da distância de Mahalanobis, respectivamente. A regressão *stepwise* para cada um dos casos com *outliers* comprovou que esses valores eram representativos das observações na amostra considerada, o que permitiu mantê-los na pesquisa.

Em relação aos pressupostos da análise multivariada, foi detectado que os indicadores apresentavam baixa linearidade e violavam a normalidade univariada e multivariada, mesmo após a utilização de escores "z". Optou-se, porém, por assumir esses resultados como limitação da pesquisa, já que dificilmente esses dois pressupostos são atendidos em estudos em Administração.

Por sua vez, os indicadores não apresentaram significativa multicolinearidade e atenderam ao pressuposto da ausência de singularidade. O teste de verificação da homogeneidade de variâncias não foi feito porque as escalas utilizadas eram ordinais com 11 categorias, ou seja, foram tratadas como intervalares. Em casos

como esse, Tabachnick e Fidell (2001) afirmam que já há a suposição de existência de uma distribuição homocedástica.

Concluída essa análise prévia, os indicadores que permaneceram associados à variável "estrutura de relacionamentos familiares e de amizade" (erfamili) foram: pai (estpai); mãe (estmae); irmão ou irmã (estirmao); outros parentes (estpare), e amigos (estamigo). Valendo-se dessa definição, foi possível construir o diagrama de caminho que representava o modelo proposto. Para realizar a AFC, foi utilizada a matriz de correlações, bem como a extração mediante a máxima verossimilhança. As estimativas padronizadas, calculadas pelo LISREL, foram as exibidas pela Figura 2.



Chi-Square = 33.21, df = 5, P - value = 0.00000, RMSEA = 0.157

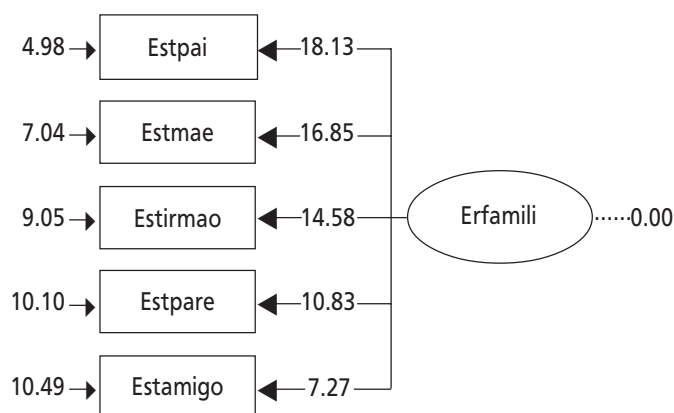
Figura 2: Estimativas padronizadas da variável estrutura de relacionamentos familiares e de amizade.

Fonte: Dados da pesquisa.

Os resultados mostrados pela Figura 2 evidenciam que o modelo é identificado, o que significa que ele foi capaz de gerar estimativas únicas. Também não foram encontradas estimativas discrepantes, posto que não havia variâncias-erro negativas, estimativas padronizadas com valor superior a 1,0 ou erros padronizados muito grandes. A ausência de estimativas discrepantes é igualmente confirmada pela análise dos valores do teste "t", exibido pela Figura 3.

Como comprova a Figura 3, todas as estimativas se situaram fora do intervalo [-2,56; 2,56]. Isso permite afirmar que todas elas puderam ser consideradas significativas em um nível de 0,01, no teste bilateral. Constatada a ausência de estimativas discrepantes, o passo seguinte foi avaliar o ajuste global do modelo. A Tabela 4 apresenta os resultados obtidos nas três medidas consideradas.

A análise do qui-quadrado e de seu valor "p" leva à rejeição da hipótese nula de que não existe diferença entre a matriz de dados e a matriz estimada, em um



Chi-Square = 33.21, df = 5, P - value = 0.00000, RMSEA = 0.157

Figura 3. Valores do teste "t" da variável "relacionamentos familiares e de amizade".

Fonte: Dados da pesquisa.

Tabela 4. Medidas de ajuste global do construto "estrutura de relacionamentos familiares e de amizade".

Medida	Valor
Qui-quadrado	33,21
Valor p	<0,00
RMSEA	0,157
Qui-quadrado normalizado	6,642

Fonte: Dados da pesquisa.

nível de significância de 0,05. Essa conclusão se fortalece quando o RMSEA é considerado. Essa medida indica que a discrepância dos dados, quando considerados os graus de liberdade, está acima do limite máximo tolerável, que é de 0,08. Por fim, o qui-quadrado normalizado mostra que, de fato, os valores estimados não foram os ideais, já que o valor encontrado nessa medida foi superior a 5,0. Portanto, considerando as medidas exibidas pela Tabela 4, pode-se afirmar que o ajuste global do modelo proposto não foi perfeito.

Em relação ao ajuste do modelo de mensuração, os resultados encontrados foram os apresentados pela Tabela 5.

Os dados da Tabela 5 indicam que há problemas no ajuste do modelo de mensuração, especialmente no que se refere aos indicadores "outros parentes" e "amigos". Os dois obtiveram, respectivamente, os valores de 0,35 e 0,26; significativamente abaixo dos demais. Por sua vez, esses também não conseguiram atingir patamares realmente expressivos, indicando que o modelo tem ajuste deficiente. No caso do exemplo adotado neste trabalho, não foi preciso recorrer à avaliação

do ajuste do modelo estrutural, uma vez que há apenas uma variável latente, que se relaciona com cinco variáveis manifestas. Todavia, se fosse necessária, a análise seria realizada da mesma maneira que a feita para o modelo de mensuração, com a única diferença de que os coeficientes de determinação expressariam as relações entre as variáveis latentes.

Tabela 5. Valores do coeficiente de determinação do constructo "estrutura de relacionamentos familiares e de amizade".

Indicador	R ²
Pai	0,73
Mãe	0,71
Irmão ou irmã	0,69
Outros parentes	0,35
Amigos	0,26

Fonte: Dados da pesquisa

Outra etapa da AFC é a avaliação da matriz de resíduos normalizados do constructo, exibida pela Tabela 6.

Tabela 6. Estrutura de relacionamentos familiares e de amizade – matriz de resíduos normalizados.

Indicador	Pai	Mãe	Irmão ou irmã	Outros parentes	Amigos
Pai	—	—	—	—	—
Mãe	1,40	—	—	—	—
Irmão ou irmã	2,89	10,21	—	—	—
Outros parentes	-0,56	2,91	3,02	—	—
Amigos	9,40	3,13	1,65	2,38	—

Fonte: Dados da pesquisa.

Como a significância adotada foi de 0,05, no máximo 5% dos dados poderiam estar fora do intervalo [-2,58; 2,58]. No entanto, a Tabela 6 mostra que 60% dos resíduos não seguiram essa regra. Com isso, mais uma vez fica demonstrado que há deficiências no ajuste do modelo proposto.

A última etapa da AFC é a verificação das estimativas padronizadas apresentadas pelo diagrama de caminho contido na Figura 2. Como é possível notar, os indicadores "pai, mãe" e "irmão ou irmã" foram os que obtiveram a relação mais forte com a variável latente "estrutura dos relacionamentos familiares e de amizade". Já os indicadores "outros parentes" e, principalmente, "amigos", apresentaram relações mais fracas com essa variável. No entanto, há que se destacar que essas conclusões devem ser consideradas com cuidado, já que o ajuste do modelo não foi o ideal.

CONCLUSÃO

A Análise Fatorial Confirmatória é uma técnica multivariada que vem sendo utilizada com relativa frequência nas pesquisas em Administração. Todavia, para que ela seja empregada corretamente, é preciso que o pesquisador tenha domínio dos aspectos conceituais e operacionais da análise multivariada.

Em primeiro lugar, antes de iniciar a AFC propriamente dita, deve-se ter a preocupação com a maneira pela qual os dados serão coletados. Para que a AFC seja conduzida com sucesso, a pesquisa precisa acessar fenômenos que possam ser expressos em relações "quantitativas". Do contrário, talvez seja mais aconselhável recorrer a outras técnicas de análise. Além da escolha do instrumento de coleta de dados, é necessário que se tenha cuidado com o tipo de escala a ser adotado na investigação. A inclusão de escalas nominais, ordinais, intervalares ou razão afeta significativamente a estratégia de pesquisa, dadas as limitações matemáticas de cada uma delas.

Por outro lado, é fundamental que o pesquisador conheça bem os seus dados antes de utilizar a AFC. A descrição do perfil da amostra adotada, bem como a análise do nível de aleatoriedade dos dados ausentes e a identificação dos valores discrepantes univariados e multivariados, é importante para que os resultados obtidos mediante uma técnica multivariada não contenham vieses significativos. Ainda com o propósito de conhecer os dados, o pesquisador não pode deixar de verificar se as informações coletadas atendem aos pressupostos da análise multivariada, em especial, a normalidade.

Depois de tomados esses cuidados iniciais, é possível proceder à AFC, cumprindo cada uma das nove etapas explicitadas neste trabalho. É preciso, porém, que o pesquisador tenha em mente que, mesmo com todo o rigor exigido nos procedimentos de uma técnica multivariada, não há garantias de que os resultados sejam isentos de problemas. Na verdade, uma dificuldade significativa na utilização da AFC nas pesquisas em Administração está no fato de que os dados raramente atendem todas as exigências para uma análise com o mínimo de vieses. Graus variados de aleatoriedade de dados ausentes, presença de valores discrepantes, violações à normalidade e baixa linearidade são alguns dos problemas mais comuns enfrentados pelos pesquisadores em Administração.

Embora essas dificuldades não inviabilizem o uso da AFC, é mister reconhecer que elas têm que ser consideradas como limitações significativas em uma pesquisa, tirando da técnica a possibilidade de ser encarada como solução universal para qualquer problema científico. Dessa forma, é aconselhável que o pesquisador planeje criteriosamente a sua investigação, ponderando os prós e os contras da estratégia de análise proposta, a fim de que a escolha feita seja realmente adequada à pesquisa.

ABSTRACT

The confirmatory factor analysis (CFA) is a statistical multivariate technique that has been frequently adopted in the investigations in Business Administration. However, the correct use of CFA demands a profound knowledge of the conceptual and operational aspects of the multivariate analysis. This paper intends to offer the researchers a guide of how to conduct a scientific investigation with the use of the confirmatory factor analysis. In order to reach this goal, some of the key theoretical aspects were mentioned in the paper. Besides that, through an empirical example, this paper called the attention to some of the main constrictions concerning to the use of CFA. The article also concludes that the inner problems of the technique are not a barrier to its use in the researches in Administration. However, it is necessary to recognize that such problems must be faced as significant limitations in the scientific investigation. Then, the researcher should seriously consider the adequacy of using the CFA in his or her investigation.

Key words: Confirmatory factor analysis; Outliers; Missing data; Multivariate analysis; Assumptions.

Referências

- AAKER, D. A. **Multivariate analysis in marketing**: theory & application. Belmont: Wadsworth, c1971. 358p.
- BISHOP, Y. M. M. **Discrete multivariate analysis**: theory and practice. Cambridge: The MIT Press, 1995.
- CHATFIELD, C.; COLLINS, A. J. **Introduction to multivariate analysis**. Boca Raton: Chapman & Hall, 1980. 246p.
- FILHO, C. G. **O Impacto da gestão do conhecimento de marketing na inovação e vantagem competitiva de novos produtos**. 2001. Tese (Doutorado em Administração) – Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Belo Horizonte.
- GOSLING, M. **Modelo estrutural de marketing de relacionamento**: aplicação e validação em uma instituição de ensino. 2004. Tese (Doutorado em Administração) – Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Belo Horizonte.
- GRAF, I. **Perceived social support versus social embeddedness**: effects of employee and organizational outcomes. 1999. 235p. Tese (Doutorado em Administração) – University of Illinois at Chicago, Chicago.
- HAIR JR., J. *et al.* **Multivariate data analysis**. 5. ed. New Jersey: Prentice Hall, 1998. 730p.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 5. ed. New Jersey: Prentice Hall, 2002. 767p.
- JÖRESKOG, K. *et al.* **LISREL 8**: new statistical features. Lincolnwood: Scientific Software International, 2000. 248 p.
- JÖRESKOG, K.; SÖRBOM, D. **LISREL 8**: user's reference guide. Lincolnwood: Scientific Software International, 2000. 378p.

- KANJI, G. K. **100 statistical tests**: new edition. London: Sage, 1999. 215p.
- LATIF, S. **Modelagem de equações estruturais**. 2000. 179p. Dissertação (Mestrado em Estatística) – Universidade de São Paulo, São Paulo.
- LOPES, H. E. G. **A força dos contatos**: um estudo das redes interpessoais de profissionais da Região Metropolitana de Belo Horizonte. 2001. 288p. Tese (Doutorado em Administração) – Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Belo Horizonte.
- LOPES, H. E. G. I know who you are: the influence of networking in business careers in Brazil. In: BALAS ANNUAL MEETING, 2003, São Paulo. **Balas 2003 Proceedings**. São Paulo: Fundação Getúlio Vargas, 2003. CD.
- MALHOTRA, N. K. **Pesquisa de marketing**: uma orientação aplicada. 3. ed. Porto Alegre: Bookman, 2001. 719p.
- MUELLER, R. **Basic principles of structural equation modeling**: an introduction to LISREL and EQS. New York: Springer, 1996. 229p. (Springer Texts in Statistics).
- MUIRHEAD, R. J. **Aspects of multivariate statistical theory**. New York: J. Wiley, c1982. [6], 673p.
- NUNNALLY, J.; BERNSTEIN, I. **Psychometric theory**. 3. ed. New York: McGraw-Hill, 1994. 752p. (McGraw-Hill Series in Psychology).
- OLSSON, U. *et al.* The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. **Structural Equation Modeling**, Hillsdale, v. 7, n. 4, p. 557-595, 2000.
- PEREIRA, J. **Análise de dados qualitativos**: estratégias metodológicas para as ciências da saúde, humanas e sociais. São Paulo: Edusp, 1999. 157p.
- SCHUMACKER, R.; BEYERLEIN, S. Confirmatory factor analysis with different correlation types and estimation methods. **Structural Equation Modeling**, Hillsdale, v. 7, n. 4, p. 629-636, 2000.
- SELLTIZ, C.; WRIGHTSMAN; COOK. **Métodos de pesquisa nas relações sociais**. 2. ed. São Paulo: EPU, 1987. 117p. (Volume 1: Delineamento de pesquisa).
- SILVA, E.; MENEZES, E. **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis: UFSC/PPGEP/LED, 2000. 118p.
- SPSS Base 10.0 applications guide. Chicago: SPSS, 1999. 426p.
- TABACHNICK, B.; FIDELL, L. **Using multivariate statistics**. 4. ed. Boston: Allyn and Bacon, 2001. 966p.
- ULLMAN, J. Structural equation modeling. In: TABACHNICK, B.; FIDELL, L. **Using multivariate statistics**. 4. ed. Boston: Allyn and Bacon, 2001. p. 653-771.
- URDAN, A.; RODRIGUES, A. O modelo do índice de satisfação do cliente norte-americano: um exame inicial no Brasil com equações estruturais. **Revista de Administração Contemporânea**, Rio de Janeiro, v. 3, n. 3, p. 109-130, set./dez. 1999.
- URDAN, A.; ZUÑIGA, M. Satisfação com assistência técnica e lealdade ao fabricante no ramo automobilístico. **Revista de Administração de Empresas**, São Paulo, v. 41, n. 3, p. 31-41, jul./set. 2001.
- VEIGA, R. **Um exame empírico do modelo de consequências comportamentais da qualidade de serviços**. 2000. 420p. Tese (Doutorado em Administração) – Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Belo Horizonte.