



Avaliação: Revista da Avaliação da Educação
Superior
ISSN: 1414-4077
revistaavaliacao@uniso.br
Universidade de Sorocaba
Brasil

Barreto De Bem, Amilton; Lanzer, Edgar Augusto; Tambosi Filho, Elmo; Sanchez, Otávio Próspero;
Bernardi Junior, Plínio

Validade e confiabilidade de instrumento de avaliação da docência sob a ótica dos modelos de
equação estrutural

Avaliação: Revista da Avaliação da Educação Superior, vol. 16, núm. 2, julio, 2011, pp. 375-401
Universidade de Sorocaba
Sorocaba, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=219119106008>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc



VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS MODELOS DE EQUAÇÃO ESTRUTURAL

AMILTON BARRETO DE BEM*

EDGAR AUGUSTO LANZER**

ELMO TAMBOSI FILHO***

OTÁVIO PRÓSPERO SANCHEZ****

PLÍNIO BERNARDI JUNIOR*****

Recebido: jun. 2010 Aprovado: dez. 2010

*Doutor em Engenharia de Produção UFSC. Professor da Universidade do Sul de Santa Catarina. Tubarão, SC, Brasil. E-mail: amilton.bem@unisul.br

**Doutor em Engenharia de Produção UFSC. Professor da Universidade do Sul de Santa Catarina. Tubarão, SC, Brasil. E-mail: edgar@unisul.br

***Doutor em Engenharia de Produção UFSC. Professor Universidade Metodista de São Paulo. São Paulo, SP, Brasil. E-mail: elmotf@hotmail.com

****Doutor em Administração de Empresas FGV/SP. Professor da Fundação Getúlio Vargas de São Paulo. São Paulo, SP, Brasil. E-mail: otavio.sanchez@metodista.br

*****Doutor em Administração de Empresas FGV/SP. Professor da Fundação Getúlio Vargas de São Paulo. São Paulo, SP, Brasil. E-mail: plinio.junior@fgv.br



Resumo: Este artigo relata um estudo sobre as propriedades de confiabilidade e validade de instrumento de avaliação da docência, utilizando técnicas de análise fatorial exploratória e confirmatória. A hipótese inicial é que o modelo de avaliação da docência da Universidade do Sul de Santa Catarina é formado por oito dimensões correlacionadas: exigência, avaliação, conteúdo, método, relacionamento professor-aluno, produtividade, ética e relacionamento professor-curso. Estudos confirmatórios revelaram, no entanto, que o modelo apresentava alguns indicadores estatísticos não adequados. Reavaliando a estrutura do instrumento por meio da técnica de análise fatorial exploratória, evidenciou-se uma estrutura com apenas quatro fatores correlacionados, quais sejam: competência, relacionamento e ética, motivação e avaliação. Submetida a testes através de técnicas de análise confirmatória, essa estrutura revelou índices de ajuste global, incremental e de parcimônia adequados sob o ponto de vista estatístico.

Palavras-chave: Instrumento Avaliação. Confiabilidade. Validade.

VALIDITY AND RELIABILITY OF THE FACULTY EVALUATION TOOLS FROM THE STRUCTURAL EQUATION MODEL PERSPECTIVE

Abstract: This article describes a study on the features concerned with the validity and reliability of the faculty evaluation tool by using corroboratory and exploratory factorial analyses techniques. The first hypothesis put forward is that the faculty evaluation model of the Universidade do Sul de Santa Catarina consists of eight correlated dimensions: requirement, evaluation, content, method, professor-student rapport, productivity, ethics and professor-course rapport. Corroboratory studies, however, have shown that the model displayed some inadequate statistical indexes. By re-evaluating the structure of the tool through an exploratory factorial analysis technique, a structure with only four correlated factors; competence, rapport and ethics, motivation and evaluation, became evident. This structure, put to the test through corroboratory analysis techniques, showed indexes of global, incremental and parsimony adjustments as being ideal from the statistical standpoint.

Key words: Tool. Evaluation. Reliability. Validity.

1 INTRODUÇÃO

Segundo Almeida (2002), o corpo docente de uma universidade é uma das principais fontes do conhecimento acadêmico. A qualidade do ensino é basicamente determinada pelas ações acadêmicas que norteiam o processo de ensino e de aprendizagem. Nesse contexto, o papel do professor pode ter especial relevância na condução da sua disciplina e na construção de um ensino mais efetivo. Assim, torna-se condição importante que uma universidade tenha um corpo docente atualizado nas diversas áreas de conhecimento e possuidor de habilidades técnicas suficientes para intermediar o processo de ensino e de aprendizagem.

Considerando que a qualidade do ensino pode também ser determinada pela capacidade do professor, a avaliação sistemática do seu desempenho, por meio de questionário com conteúdo técnico adequado, conduz os professores, coordenadores e outros agentes à análise crítica e positiva dos resultados decorrentes do processo de avaliação.

Aubrecht (1979) diz que a qualidade das conclusões de um processo de avaliação sobre a qualidade do ensino depende da existência de um conjunto de dimensões confiáveis e válidas. Já Moreira (1986), mostra que um instrumento bem elaborado é condição necessária, embora não suficiente, para o sucesso de uma pesquisa de avaliação.

Outro aspecto importante para o sucesso da pesquisa é assegurado quando se utiliza indicativos de elevada consistência interna entre os alunos que estão freqüentando determinada disciplina e que há evidência empírica de que cada dimensão do instrumento reflete uma característica importante da qualidade do ensino, ficam evidenciadas as propriedades fundamentais do instrumento de medida, a confiabilidade e validade. Tais propriedades estatísticas, embora mencionadas nos diversos conceitos de avaliação, nem sempre são discutidas cientificamente. Esse fato é facilmente comprovado pelo número insignificante de publicações que tratam sobre o tema na literatura brasileira.

Malhotra (2001) define a confiabilidade como o grau em que uma escala produz resultados consistentes quando se realizam medições repetidas das características em análise. Para Ramos (1987) a validade de um instrumento de medida é o grau com que os indicadores unidimensionais da escala medem com consistência interna as dimensões que foram designadas para medir.

Sob esta ótica, esse artigo investiga a importância dos dados empíricos provenientes do processo de avaliação de disciplinas de uma universidade, com o uso de técnicas de análise fatorial exploratória e confirmatória, para



VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS MODELOS DE EQUAÇÃO ESTRUTURAL

testar as propriedades de confiabilidade e validade das dimensões subjacentes no instrumento de avaliação dos docentes pelos discentes.

2 ASPECTOS TEÓRICOS DA CONFIABILIDADE E DA VALIDADE DAS AVALIAÇÕES DOS DOCENTES PELOS DISCENTES

Pedhazur e Schmelkin (1991) mostram que uma medida é confiável e válida quando associada com o menor erro possível ao conceito que está sendo medido. Há três categorias de erros que podem estar ligados a uma pesquisa: erros relacionados ao pesquisador, os erros relacionados ao instrumento e aqueles relacionados ao respondente. Os erros de medida relacionados ao instrumento podem ser reduzidos, principalmente pela construção de instrumento de medida bem planejado. Uma solução é optar por medições multivariadas, em que diversos itens da escala inter-relacionados proporcionam uma medida composta do conceito operacionalizado. Segundo Hair et al (1999), o uso de vários indicadores inter-relacionados proporciona uma medição do conceito numa perspectiva mais completa.

O coeficiente α , proposto por Cronbach (1951), para quantificar a confiabilidade de instrumentos de medidas multidimensionais, considera a homogeneidade dos itens da escala e apresenta como vantagem o fato de necessitar de uma única aplicação do instrumento. É o método mais utilizado para medir a confiabilidade, quando esta for entendida como uma consistência interna dos indicadores da escala, ou seja, os indicadores da escala, altamente inter-relacionados, devem medir o mesmo construto latente.

Confiabilidade e validade são duas propriedades de medida estreitamente relacionadas, desempenhando papéis complementares. A confiabilidade de consistência interna está relacionada à homogeneidade das respostas dos distintos avaliadores, enquanto a validade está associada ao grau de certeza que se tem sobre o conceito medido.

Para Ramos (1987), o tema validez é complexo devido ao caráter indireto da medição, já que nunca se mede uma variável latente. O que se mede são algumas manifestações do domínio de condutas dos indivíduos. Nesse sentido, fica difícil manifestar se o instrumento é válido ou não válido e, em que grau é válido para medir o que se pretende.

Segundo Hair Jr. et al (2005) o processo de analisar a validade de um instrumento depende, dentre outros fatores, das variáveis a validar, dos objetivos do instrumento de medida e da população a ser submetida. A literatura discute

várias formas de se garantir a validade de um instrumento: validade de conteúdo, de construto, convergente, discriminante. A validação de conteúdo é feita por meio do julgamento do pesquisador ou de especialistas quanto ao conteúdo do instrumento. A validade do construto procura avaliar se a escala está medindo, de fato, o que se propõe a medir. Ela pode ser conseguida por meio de técnicas estatísticas, como propõe essa pesquisa. A validade convergente mede a coerência e a uniformidade entre indivíduos semelhantes. A validade discriminante verifica o ponto até onde o construto não se correlaciona com outros construtos que dele diferem.

Segundo Tourón (1989), quando se valida um construto, se valida também o conjunto de teorias nas quais ele está inserido. Portanto, é toda uma rede de relações que é submetida à prova. A validade de construto engloba, pois, os tipos de validade. Com suas múltiplas técnicas e procedimentos, é a trajetória adequada para a validação das dimensões subjacentes no instrumento.

Nunnally (1978) descreve um processo de três passos para providenciar a validade de construto de um instrumento: 1) especificar o domínio de indicadores inerentes ao construto; 2) a partir de investigação empírica e análises estatísticas, determinar até que ponto os indicadores tendem a medir um único construto ou diversos construtos; e 3) conduzir estudos diferenciais e/ou experimentos controlados para determinar até que ponto os indicadores supostos do construto produzem resultados que são previsíveis a partir de hipóteses teóricas aceitas sobre o construto.

Loesch e Hoeltgebaum (2005) afirmam que os modelos de equações estruturais são particularmente úteis nas ciências sociais e do comportamento, sendo, portanto, adequados para testar a estrutura fatorial de instrumento de avaliação da docência.

Para Ramos (1987), o processo de validação de construto é um processo iterativo, interativo e de integração de resultados. Iterativo porque requer a repetição de um processo esquemático básico para confirmação de hipóteses relativas ao construto. Interativo porque exige estudos diversificados que permitam estabelecer as covariâncias entre as variáveis relativas ao construto. Integração de resultados porque é necessário que os distintos resultados sobre a natureza do construto, sejam interpretados de forma integrada, descobrindo que contradições se manifestam entre os resultados dos distintos estudos e quais as suas coincidências.

Ainda, segundo Ramos (1987), o modelo de análise fatorial confirmatória é muito útil em processos de validação de construto, especialmente quando se pretende validar determinada hipótese estrutural decorrente da análise lógica do conteúdo ou validar outras hipóteses estruturais alternativas.



3 AVALIAÇÕES DAS PROPRIEDADES DO INSTRUMENTO DE AVALIAÇÃO MEDIANTE MODELOS DE ANÁLISE FATORIAL EXPLORATÓRIA E CONFIRMATÓRIA

Comumente, à análise factorial exploratória é utilizada para analisar os inter-relacionamentos lineares dentro de um conjunto de indicadores observados, contados ou medidos para cada indivíduo de um grupo, postulando que, um conjunto reduzido de fatores comuns é suficiente para explicar os inter-relacionamentos existentes entre os indicadores. Devido a sua função exploratória não assume nenhum modelo de medida. Sua única finalidade é a de reduzir a dimensionalidade. Também não existe uma cota mínima para a variância total explicada pela estrutura fatorial. Portanto, diferentes critérios podem ser utilizados como critério de corte. A diversidade de métodos de rotação pode conduzir a interpretações distintas.

Em face do exposto, dotar determinado instrumento de avaliação de certo grau de validade de construto com base na estrutura fatorial decorrente da análise factorial exploratória é insuficiente, pois ela nem sempre reflete a realidade dos fatos, necessitando, por isso, de outras técnicas com, por exemplo, análise factorial confirmatória.

Segundo Jöreskog (1969), o modelo de análise factorial confirmatória, caso particular dos modelos de equação estrutural, corrige as deficiências do modelo exploratório e conduz a uma maior certeza das hipóteses que devem ser contrastadas através de modelos que expliquem todos os inter-relacionamentos existentes na estrutura de um questionário.

Os modelos de equação estrutural com variáveis latentes e erros de medida utilizados como estratégia de modelização confirmatória por incluído Marsh (1991), Mateo e Fernández (1992), Ramos (1997), Val (1994), Greimel-Fuhrmann e Geyger (2003), incluído Capelleras e Veciana (2000), comprovam a versatilidade dos modelos confirmatórios para fornecer o suporte empírico de validade de construto de instrumento de avaliação da docência.

Hair et al (1999) propõem as seguintes fases para testar as hipóteses levantadas sobre a multidimensionalidade de instrumento: a especificação do modelo; a identificação do modelo; a estimativa do modelo; e a avaliação e interpretação do modelo proposto.

O diagnóstico de bondade do ajuste do modelo estimado é crucial para o estabelecimento da validade de construto. Por isso, diversos testes de diagnóstico e índices foram desenvolvidos. Destacam-se as medidas de ajuste absoluto, as medidas incrementais de ajuste e as medidas de ajuste de parcimônia.



AMILTON BARRETO DE BEM; EDGAR AUGUSTO LANZER; ELMO TAMBOSI FILHO;
OTÁVIO PRÓSPERO SANCHEZ; PLÍNIO BERNARDI JUNIOR

As medidas de ajuste absoluto determinam o grau em que o modelo de medida construído é capaz de predizer com o menor erro possível a matriz de variância-covariância ou a matriz de correlação utilizada na modelagem. Destaca-se o teste de χ^2 (Qui-Quadrado) citado por Marsh, Balla e McDonald (1988), a medida de parâmetro de não centralidade (PNC) proposta por McDonald e Marsh (1990), o índice de adequação do ajuste (GFI) proposto por Jöreskog e Sörbom (1984 apud MULAIK et al, 1989) e o índice quadrático médio (RMSR) proposto Jöreskog e Sorbom (1984 apud MARSH et al, 1988).

Medidas incrementais de ajuste são indicadores da qualidade, que possibilitam comparar modelos distintos. Nessa categoria tem-se o índice de adequação corrigido (AGFI) proposto por Jörekog e Sorbom (1981 apud MARSH et al, 1988), o índice Tucker-Lewis (TLI) proposto por Tucker e Lewis (1973) e o índice de ajuste normal (BBI), proposto por Bentler e Bonnett (1980).

Um modelo é parcimonioso quando possui boa capacidade explicativa e não contém coeficientes desnecessários. As medidas de ajuste de parcimônia relacionam a bondade do modelo com o número de coeficientes necessários para alcançar esse nível de ajuste. Destaca-se aqui o índice de ajuste normalizado de parcimônia (PNFI), proposto por MulaiK et al (1989); índice de bondade de ajuste de parcimônia (PGFI) também proposto por MulaiK et al (1989).

Greimel-Fuhrmann, e Geyger (2003) recomendam que, embora tenhamos bons indicadores do ajuste global do modelo de medida estimado, seja feita uma análise individual dos indicadores exógenos ou endógenos para evitar parâmetros estimados não significativos. Assim, se alguns dos indicadores do modelo de medida não forem estatisticamente significativos, conforme indica o teste de contraste t student, o indicador deve ser eliminado ou transformado para que se ajuste melhor ao construto.

Segundo Ramos (1986), a validade de um instrumento é uma questão de grau, uma vez que em todo processo de estimativa, efetua-se inferência sobre os verdadeiros parâmetros populacionais com um dado nível de significância. Assim, nunca será possível afirmar que um modelo de medida seja totalmente correto e tampouco demonstrar que seja totalmente incorreto.

4 APLICAÇÃO DAS TÉCNICAS DE ANÁLISE FATORIAL EXPLORATÓRIA E CONFIRMATÓRIA AO INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA DA UNIVERSIDADE DO SUL DE SANTA CATARINA

Esse estudo fez uso de instrumento de avaliação docente aplicado pela Universidade do Sul de Santa Catarina (UNISUL) para a investigação das pro-



VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS MODELOS DE EQUAÇÃO ESTRUTURAL

priedades psicométricas de confiabilidade e validade do instrumento “Avaliação pelos Alunos do Desempenho na Disciplina”, composto de oito dimensões e vinte e nove indicadores.

Esta pesquisa tem característica explicativa, já que propõe demonstrar a validade discriminante dos indicadores e construtos teóricos. Do ponto de vista da forma de abordagem do problema é quali-quantitativa. O método é o hipotético-dedutivo.

Os indicativos de confiabilidade e de validade dependem de que as avaliações realizadas estejam isentas de erro aleatório e/ou sistemático. Uma forma de minimizar o erro aleatório é adotar a média da turma em cada um desses indicadores como unidade de análise. Segundo Hair et al, (1999), o uso das médias como unidade de análise atenua o erro de medida inerente a todos os indicadores medidos, uma vez que os erros individuais de estimativa tendem a se anular e a média aparece como um indicador puro. Marsh (1988), também argumenta que a resposta média da classe de alunos, como unidade de análise, equivale a um acordo entre todos os avaliadores de uma mesma classe e deve ser utilizada para acessar a confiabilidade e a validade das dimensões da qualidade do ensino medidas pelo instrumento de avaliação.



A aplicação das técnicas multivariadas de análise fatorial exploratória e confirmatória implica que os indicadores analisados cumpram as suposições estatísticas de normalidade, de linearidade e de homocedasticidade. A garantia de linearidade dos indicadores é o que possibilita o uso da análise fatorial exploratória com vistas à obtenção de uma estrutura simplificada, enquanto que a violação dos pressupostos de normalidade e homocedasticidade reduzem o poder dos testes paramétricos utilizados para avaliar as hipóteses estabelecidas. Assim, a violação desses pressupostos, por parte de algum indicador medido, indica que eles devem ser transformados por técnicas especiais.

O modelo qualitativo teórico, apresentado no diagrama de caminhos 1, constitui a primeira hipótese a ser testada e sustenta a afirmação de que o instrumento é composto por oito dimensões latentes, confiáveis e correlacionadas.

Na determinação do tamanho das amostras é levada em consideração a sensibilidade do método de estimativa de máxima verossimilhança para grandes amostras e o fato de que os dados empíricos, mesmo transformados, não apresentem coeficiente de normalidade multivariada de Mardia igual à zero. Ponderando sobre os critérios acima, foi utilizada inicialmente uma amostra de 675 disciplinas escolhidas aleatoriamente do total de 2042 disciplinas avaliadas.

Foi utilizada a matriz de correlação para a estimativa dos parâmetros, pelo método de estimativa de máxima verossimilhança (MV). Esse método foi escolhido pelas boas propriedades assintóticas dos estimadores (BOLLEN, 1989).



AMILTON BARRETO DE BEM; EDGAR AUGUSTO LANZER; ELMO TAMBOSI FILHO;
OTÁVIO PRÓSPERO SANCHEZ; PLÍNIO BERNARDI JUNIOR

5 RESULTADOS E DISCUSSÃO

A análise fatorial exploratória e confirmatória requer que as distribuições das medidas utilizadas cumpram os critérios de homocedasticidade, linearidade e normalidade multivariada.

Quanto à normalidade dos dados empíricos, o coeficiente de normalidade multivariada de Mardia (1989) foi igual a 0,450 e após a transformação dos dados alcançou 0,180. A transformação dos dados empíricos, além de diminuir o excesso de assimetria e de curtose conduz a redução da variabilidade dos indicadores medidos.

Durante a fase de especificação do modelo de medida de indicadores exógenos, deve-se considerar a confiabilidade dos indicadores e dos construtos. Todos os indicadores desta pesquisa apresentaram correlação item-total superiores a 0,5 e dimensões com confiabilidade superior a 0,7, comumente evidenciada na literatura. Esses resultados confirmam que as oito dimensões do instrumento de avaliação de disciplina da UNISUL são consistentemente avaliadas pelos discentes. Indicam, ainda, que o modelo teórico especificado no diagrama 1 é testado a partir de medidas com erros aleatórios minimizados.

Os valores dos parâmetros lâmbda (λ) e o seu respectivo erro padrão, o erro de medida theta delta (δ) dos indicadores decorrentes do processo de estimativa por máxima verossimilhança do modelo com oito fatores correlacionados seguem apresentados na tabela 1.

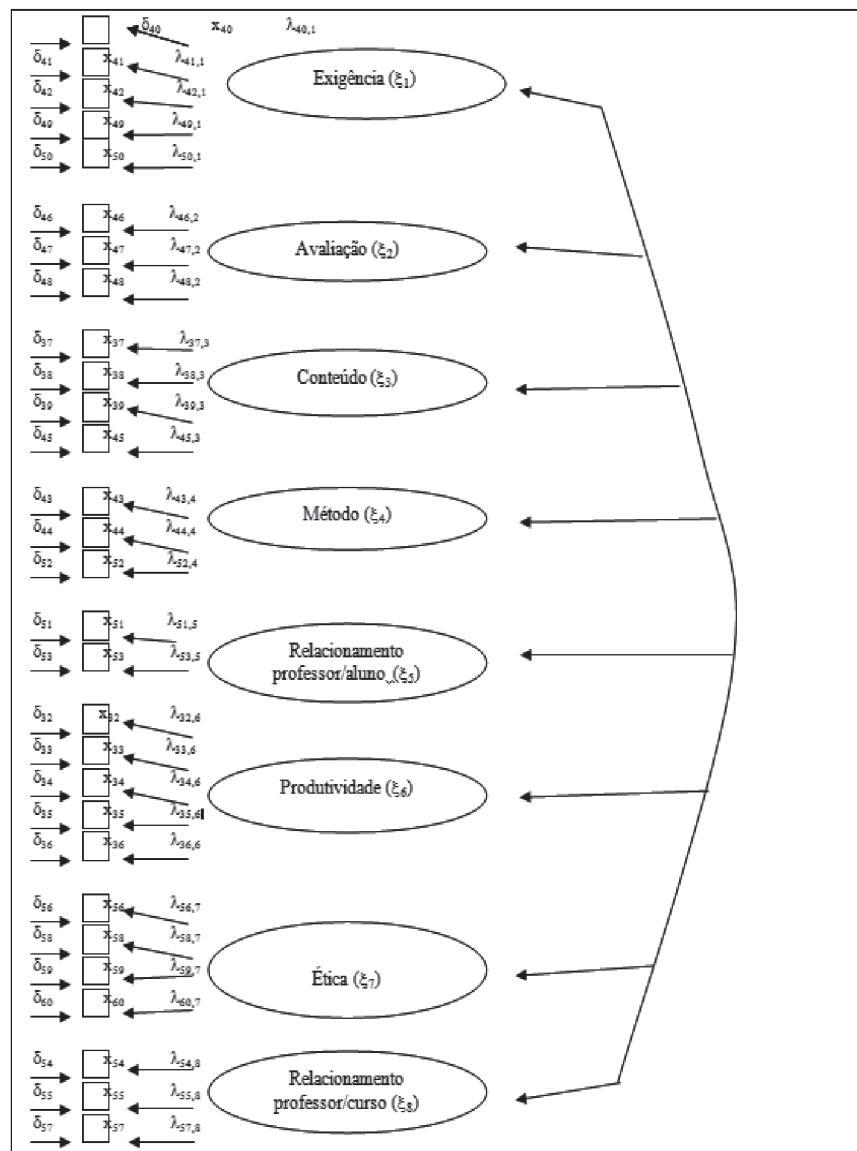
O modelo de medida de indicadores exógenos, com oito fatores correlacionados, apresenta todos os coeficientes (λ) positivos e estatisticamente significativos. Isso pode ser constatado pelo baixo erro padrão evidenciado pelos parâmetros estimados. Entretanto, os indicadores de números 41 e 42 apresentam elevado erro de medida (theta delta).

O gráfico de probabilidade normal para os resíduos do modelo com oito fatores correlacionados, apresentado no gráfico 1, exibe alguns resíduos padronizados que superam $\pm 2,58$, sugerindo que o modelo proposto não se encaixa bem aos dados empíricos, podendo gerar como consequência erro de aproximação quadrático médio alto.



VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

**Diagrama 1 - Diagrama de caminhos das características
da qualidade docente (oito fatores latentes correlacionados)**



Fonte: Avaliação Institucional da UNISUL (1995).

AMILTON BARRETO DE BEM; EDGAR AUGUSTO LANZER; ELMO TAMBOSI FILHO;
OTÁVIO PRÓSPERO SANCHEZ; PLÍNIO BERNARDI JUNIOR

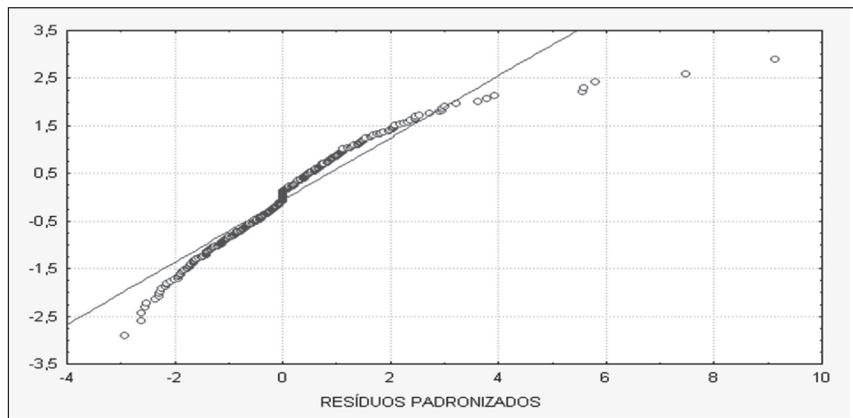
**Tabela 1 - Valores dos coeficientes lâmbda (λ),
erro padrão e erro de medida dos indicadores**

Dimensão	Máxima verossimilhança		
	Lâmbda (λ)	Erro padrão	Theta delta (δ)
Exigência (1)			
40	0,7106	0,020	0,495
41	0,5703	0,027	0,675
42	0,5901	0,026	0,652
49	0,890	0,009	0,207
50	0,9432	0,005	0,110
Avaliação (2)			
46	0,886	0,010	0,215
47	0,855	0,012	0,268
48	0,935	0,007	0,125
Conteúdo (3)			
37	0,829	0,012	0,314
38	0,834	0,012	0,304
39	0,898	0,008	0,193
45	0,822	0,013	0,323
Método (4)			
43	0,885	0,010	0,216
44	0,816	0,013	0,334
52	0,853	0,011	0,272
Relacionamento professor aluno (5)			
51	0,871	0,011	0,241
53	0,860	0,012	0,261
Produtividade (6)			
34	0,943	0,010	0,110
36	0,748	0,017	0,441
Ética (7)			
56	0,814	0,014	0,337
58	0,827	0,013	0,317
59	0,899	0,008	0,192
60	0,891	0,009	0,206
Relação professor curso (8)			
54	0,884	0,011	0,219
55	0,78	0,011	0,230
57	0,797	0,016	0,365

Fonte: Elaborada pelos autores (2005).

Notas: (1) Os indicadores de números 32, 33 e 35 foram excluídos da análise em função de evidenciarem baixas variabilidades, consequentemente, evidenciam pouca contribuição na formação deste construto. (2) Todos os lâmbdas (λ) são estatisticamente significativos ao nível de 1%.

**Gráfico 1 - Gráfico de probabilidade normal dos resíduos do
modelo com oito fatores correlacionados, geral UNISUL**



Fonte: Elaborado pelos autores (2005).

A tabela 2 apresenta as medidas de adequação do ajuste do modelo com oito fatores latentes correlacionados e do modelo com oito fatores latentes não correlacionados.

Quanto aos índices de ajuste absoluto, o modelo com oito fatores correlacionados é o que apresenta melhores resultados. O índice RMSR diminui em 45% quando se compara o modelo com oito fatores correlacionados ao modelo com oito fatores independentes. Ressalta-se que o modelo com oito fatores correlacionados apresenta erro de aproximação quadrático médio muito acima do recomendado, Browne e Cudec (1993) sugeriram que valores de RMSEA menores do que 0,05 são indicadores de bom ajuste. Um ajuste perfeito corresponde a RMSEA=0,00. Segundo Aubrecht (1979), ou seja, acima de 0,08. Os demais índices de ajuste absoluto também são menores no modelo com oito fatores correlacionados.

Em relação aos índices incrementais de ajuste, os índices AGFI, TLI e BBI são mais elevados no modelo com oito fatores correlacionados, indicando que este apresenta maior poder de explicação quando comparado com o modelo de oito fatores independentes. Ainda assim, o modelo com oito fatores correlacionados precisa ser melhorado para alcançar melhor poder de explicação.

Os índices de ajuste de parcimônia PNFI, PGFI e o AIC confirmam o modelo com oito fatores correlacionados como superior ao modelo com oito fatores independentes.

Quanto à correlação entre os fatores latentes, constatou-se que os fatores 3 e 1, 4e1, 6e1, 4e3, 6e3, 5e4 e 7e5 apresentam correlações superiores a 0,9. Assim sendo, não evidenciam construtos diferentes. Quanto à confiabilidade composta e à variância extraída desses fatores latentes, todos os resultados superam os limites mínimos de aceitação, ou seja, 0,7 para a confiabilidade composta e 0,5 para a variância extraída.

Tabela 2 - Índices de adequação do ajuste dos modelos com oito fatores não correlacionados e oito fatores correlacionados

Medidas de bondade do ajuste	Geral da UNISUL	
	Modelo com oito fatores não correlacionados	Modelo com oito fatores Correlacionados
Absoluto de ajuste		
χ^2	10283,40 (gl 301)	3474,20 (gl 271)
RMSR	0,269	0,149
GFI	0,370	0,668
PNC	10522,4	3203,20
PNCE	15,589	4,745
Incrementais de ajuste		
AGFI	0,265	0,571
TLI	0,471	0,812
BBI	0,503	0,832
Ajuste de parcimônia		
PNFI	0,466	0,694
PGFI	0,3427	0,5571
χ^2 (Normalizado)	34,164	12,820
AIC	15,406	5,392

Fonte: Elaborada pelos autores (2005).

Notas: (1) χ^2 : Estatística Qui-quadrado ($p_{valor} > 0,05$), (2) RMSR: Erro de aproximação quadrático médio ($< 0,08$), (3) RMSR: Raiz quadrada média dos resíduos (próximo de zero), (4) GFI: Índice de adequação do ajuste ($> 0,90$), (5) PNC: Parâmetro de não centralidade (menor melhor), (6) PNCE: Parâmetro de não centralidade da escala (menor melhor), (7) AGFI: Índice de adequação do ajuste corrigido ($> 0,90$), (8) TLI: Índice de ajuste de Tucker-Lewis ($> 0,90$), (9) BBI: Índice de ajuste normalizado ($> 0,90$), (10) PNFI: Índice de ajuste normalizado de parcimônia (maior melhor), (11) PGFI: Índice de bondade de ajuste (maior melhor), (12) AIC: Critério de informação de Akaike (menor melhor).

Embora, todos os indicadores de confiabilidade assegurem que os modelos foram estimados, utilizando-se de medidas confiáveis, os índices de adequação do ajuste absoluto, incremental e de parcimônia não são satisfatórios sob o ponto de vista estatístico. Existem indicadores que apresentam elevados erros de medida; há fatores latentes que são altamente correlacionados e uma quantidade de resíduos que excede ao valor de corte resultante do nível de significância escolhido. Isso remete a uma reavaliação da estrutura fatorial.

O estudo das dimensões é realizado mediante a aplicação dos dados empíricos à análise fatorial exploratória. A análise da matriz de correlação entre todos os indicadores da escala através do teste de esfericidade de Bartlett e do índice de Kaiser-Meyer-Olkin, confirmam que os indicadores estão correlacionados e que suas características são adequadas para realizar uma análise fatorial exploratória.

Os números de fatores a serem retidos foram evidenciados pelo critério de contraste de caída. O procedimento de rotação ortogonal varimax, que é um método de rotação para cada componente principal com alguns pesos significativos com todos os outros próximos de zero, tendo como função maximizar a variação entre os pesos de cada componente principal. Foi o que melhor evidenciou a estrutura desses fatores. Posteriormente, realizou-se o procedimento de rotação oblíqua, com o objetivo de determinar os níveis de correlação entre os fatores. Os níveis de saturação entre os indicadores e os fatores retidos mantiveram-se os mesmos, nos dois tipos de rotação.

A tabela 3 apresenta a composição dos fatores, a communalidade de cada indicador; a matriz de correlação entre os fatores e confiabilidade destes; o autovalor; a porcentagem da variância explicada e a porcentagem de variância explicada acumulada.

Ressalta-se que os indicadores x_{38} , x_{43} , x_{44} , x_{45} , x_{54} , x_{55} , x_{56} , x_{57} e x_{59} foram removidos da análise, por não apresentarem carga fatorial suficiente ou por não se discriminarem bem entre os fatores. A cada uma dessas dimensões atribuiu-se um significado conforme se vê abaixo.

- **Fator 1:** Competência que reconhece que o professor que administra o tempo em sala de aula de forma produtiva; desenvolve o plano de ensino; revela domínio de conteúdo na sua disciplina; apresenta uma bagagem intelectual estimulante para os alunos; trabalha a disciplina com um nível de exigência suficiente e conduz o ensino da disciplina com qualidade, possui os atributos mais importantes para uma docência com qualidade. Essa constatação decorre do fato de que o primeiro fator retido pela análise fatorial é o que carrega a maior porcentagem da variância explicada. Engloba cinco indicadores impor-

tantes que enfatizam a capacidade do professor em planejar e organizar a sua disciplina. Esse fator, segundo a meta-análise realizada por Cohen (1981), está presente na maioria dos instrumentos de avaliação da docência. Devido à característica generalista manifestada, esse fator é rotulado como “competência” e é relacionado positivamente com os demais fatores.

- **Fator 2:** Relacionamento que diz respeito ao professor. Engloba os seguintes atributos que se fazem necessários a uma boa docência, quais sejam: o bom relacionamento com os alunos, o valor dado à participação dos alunos nas aulas, estímulo ao bom relacionamento da classe de alunos, o respeito pelo aluno como pessoa e a valorização da cooperação e da solidariedade em sala de aula. Este fator é rotulado como “relacionamento e ética”. É o segundo fator mais importante, e presente na maioria dos instrumentos de avaliação da docência.

- **Fator 3:** Exigência que estão relacionados nesse fator, os indicadores que visam a um ensino mais efetivo. São eles: estímulo à utilização de bibliografia atualizada, a produção científica dos alunos e incentivo à leitura de livros, textos, jornais e revistas complementares às aulas. Esse fator está teoricamente correlacionado com o primeiro e o segundo fator e é denominado de “motivação dos discentes”.

- **Fator 4:** Avaliação que mostra que os indicadores relacionados com esse fator são: o professor redige provas ou verificações de aprendizagem de forma clara; informa sobre os critérios adotados na avaliação; e planeja as avaliações compatíveis com os objetivos e conteúdos ministrados. Esse fator denomina-se “avaliação”. Está presente em todos os instrumentos de avaliação da docência, e correlaciona-se com os demais.

Portanto, pode-se dizer que o instrumento de avaliação da docência pelos alunos que, *a priori*, foi concebido com oito dimensões, pode ser explicado por quatro dimensões ou fatores correlacionadas que explicam aproximadamente 83% da variabilidade dos indicadores retidos para análise, quais sejam: competência, relacionamento e ética, motivação, e avaliação. Os indicadores dessas dimensões são unidimensionais, evidenciando a sua validade discriminante.

Tendo assegurado que os fatores são adequados às suas definições conceituais, que os indicadores são unidimensionais e que alcançam níveis suficientes de confiabilidade, faz-se necessária uma avaliação final, ou seja, a validação da escala através de modelos de análise fatorial confirmatória. Testaram-se duas hipóteses. A primeira sustenta que todos os indicadores retidos pela análise fatorial estão associados a um único construto teórico (modelo nulo). A Segunda, testa a hipótese de que os quatro fatores correlacionados retidos pela análise fatorial exploratória são confiáveis e válidos.

VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

Tabela 3 - Matriz fatorial rotacionada

Indicadores	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidade
X ₃₄	0,8380				0,7846
X ₃₆	0,6816				0,5721
X ₃₇	0,8148				0,7369
X ₃₉	0,7445				0,8083
X ₄₉	0,7941				0,8035
X ₅₀	0,7783				0,8766
X ₅₁		0,8146			0,8117
X ₅₂		0,7003			0,7303
X ₅₃		0,8111			0,7745
X ₅₈		0,8232			0,7470
X ₆₀		0,7917			0,7922
X ₄₀			0,7268		0,6805
X ₄₁			0,8381		0,5768
X ₄₂			0,8679		0,6690
X ₄₆				0,7426	0,7342
X ₄₇				0,7207	0,6947
X ₄₈				0,7035	0,7881
Correlações entre fatores					
	Fator 1	Fator 2	Fator 3	Fator 4	
Fator 1	0,9491				
Fator 2	0,6917	0,9406			
Fator 3	0,5500	0,6230	0,8752		
Fator 4	0,7681	0,7479	0,5037	0,9184	
Autovalor	10,4780	1,5738	1,2977	0,6862	
% Variância explicada	61,6366	9,2574	7,6337	4,0366	
% Variância explicada acumulada	61,6366	70,8940	78,5277	82,5643	

Fonte: Elaborada pelos autores (2005).

Nota: Elementos em negrito na diagonal principal da matriz de correlação entre os fatores correspondem à confiabilidade do fator.

Submetendo a matriz de correlação dos indicadores retidos pela análise factorial exploratória ao método de estimação por máxima verossimilhança (ML), têm-se os índices de qualidade do ajuste apresentados na tabela 4.

Quanto aos índices de ajuste absoluto, o modelo com quatro fatores correlacionados é o que apresenta melhor resultado em todos os tipos de índice. O valor do erro de aproximação quadrático médio (RMSER) é de 0,094, podendo estender-se entre 0,084 e 0,104, com nível de confiança de 90%. A raiz quadrada média dos resíduos é de 0,044. O gráfico de probabilidade normal para os resíduos padronizados, apresentados no gráfico 2, exibe todos os resíduos padronizados no intervalo de -3 a +3, o que indica que a matriz de correlações previstas não se distancia da matriz de correlação observada, com confiança estatística de 99,73%.

O índice GFI indica que 87,3% da variabilidade dos indicadores saturados pelos seus respectivos construtos são por ele explicados. Quanto aos índices PNC e PNCE, esses também são maiores no modelo com quatro fatores correlacionados. Do exposto conclui-se que o modelo com quatro fatores correlacionados, prevê satisfatoriamente a matriz observada.

Em relação aos índices incrementais de ajuste, também o modelo com quatro fatores correlacionados é o que obtém melhores resultados. Os índices TLI e BBI, respectivamente iguais a 0,933 e 0,938 são superiores ao do modelo nulo. Em síntese, esses dois índices superam a cota mínima necessária para julgar adequados os construtos operacionalizados no instrumento e indicam que o modelo com quatro fatores correlacionados é superior ao modelo nulo.

Os índices de ajuste de parcimônia confirmam o modelo com quatro fatores correlacionados como o mais adequado para reproduzir a matriz de correlação amostral. Destaca-se, nesse bloco, o critério de informação de *Akaike*, adequado para comparar modelos com diferentes números de construtos. Valor próximo de zero indica que o modelo estimado é parcimonioso. Todos esses índices visualizados na tabela 4, indicam que os dados da amostra se ajustam satisfatoriamente ao modelo proposto de quatro fatores latentes correlacionados.

A análise detalhada dos fatores latentes é realizada examinando-se a significância de cada indicador, os erros de medida, a confiabilidade dos indicadores e dos fatores latentes. Todos esses indicadores da adequação do modelo estão apresentados na tabela 5. Na parte inferior da tabela apresenta-se a matriz de correlação entre os quatro fatores latentes com os índices de confiabilidade composta na diagonal principal.

VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

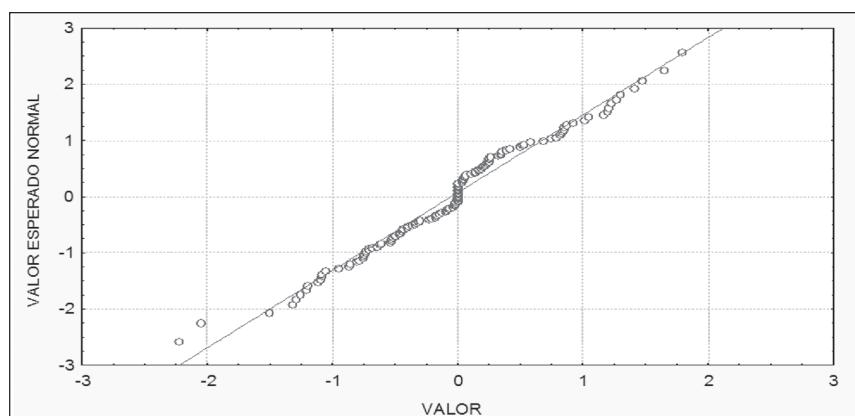
**Tabela 4 - Índices de adequação do ajuste do
modelo nulo e do modelo com quatro fatores**

Medidas de bondade do ajuste	Geral da UNISUL	
	Modelo com único fator latente	Modelo com 4 fatores latentes correlacionados
Absoluto de ajuste		
χ^2	1228,83 (gl 104)	375,235 (gl 98)
RMSER	0,264	0,094 (0,084; 0,104)
RMSR	0,093	0,0445
GFI	0,511	0,873
PNC	1124,83	277,235
PNCE	4,687	0,840
Incrementais de ajuste		
AGFI	0,360	0,824
TLI	0,691	0,933
BBI	0,715	0,938
Ajuste de parcimônia		
PNFI	0,620	0,762
PGFI	0,443	0,713
χ^2 (Normalizado)	11,816	3,829
AIC	5,409	1,382

Fonte: Elaborada pelos autores (2005).

Notas: (1) χ^2 : Estatística Qui-quadrado ($p_{valor} > 0,05$), (2) RMSER: Erro de aproximação quadrático médio ($< 0,08$), (3) RMSR: Raiz quadrada média dos resíduos (próximo de zero), (4) GFI: Índice de adequação do ajuste ($> 0,90$), (5) PNC: Parâmetro de não centralidade (menor melhor), (6) PNCE: Parâmetro de não centralidade da escala (menor melhor), (7) AGFI: Índice de adequação do ajuste corrigido ($> 0,90$), (8) TLI: Índice de ajuste de Tucker-Lewis ($> 0,90$), (9) BBI: Índice de ajuste normalizado ($> 0,90$), (10) PNFI: Índice de ajuste normalizado de parcimônia (maior melhor), (11) PGFI: Índice de bondade de ajuste (maior melhor), (12) AIC: Critério de informação de Akaike (menor melhor).

**Gráfico 2 - Gráfico de probabilidade normal para os
resíduos do modelo com quatro fatores correlacionados**



Fonte: Elaborada pelos autores (2005).

Tabela 5 - Análise factorial confirmatória. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – Geral UNISUL

Fator	Lâmbda (λ^*)	Erro padrão	Theta delta (δ)	Confiabilidade do indicador
Competência (ξ_1)				
X_{34}	0,879	0,014	0,227	0,773
X_{37}	0,859	0,016	0,262	0,738
X_{39}	0,897	0,012	0,195	0,805
X_{49}	0,896	0,012	0,198	0,803
X_{50}	0,959	0,006	0,080	0,920
	0,808**			
Relacionamento/ ética (ξ_2)				
X_{51}	0,868	0,015	0,246	0,753
X_{52}	0,875	0,015	0,234	0,766
X_{53}	0,877	0,015	0,231	0,769
X_{58}	0,844	0,018	0,287	0,712
X_{60}	0,908	0,012	0,176	0,824
	0,765**			
Exigência (ξ_3)				
X_{40}	0,861	0,021	0,258	0,741
X_{41}	0,786	0,026	0,382	0,618
X_{42}	0,852	0,021	0,274	0,726
	0,695**			
Avaliação (ξ_4)				
X_{46}	0,886	0,014	0,215	0,785
X_{47}	0,854	0,017	0,270	0,729
X_{48}	0,936	0,011	0,124	0,877
	0,797**			
Valores phi				
ξ_1	0,954	ξ_2	ξ_3	ξ_4
ξ_2	0,768	0,942		
ξ_3	0,689	0,637	0,872	
ξ_4	0,803	0,820	0,581	0,922

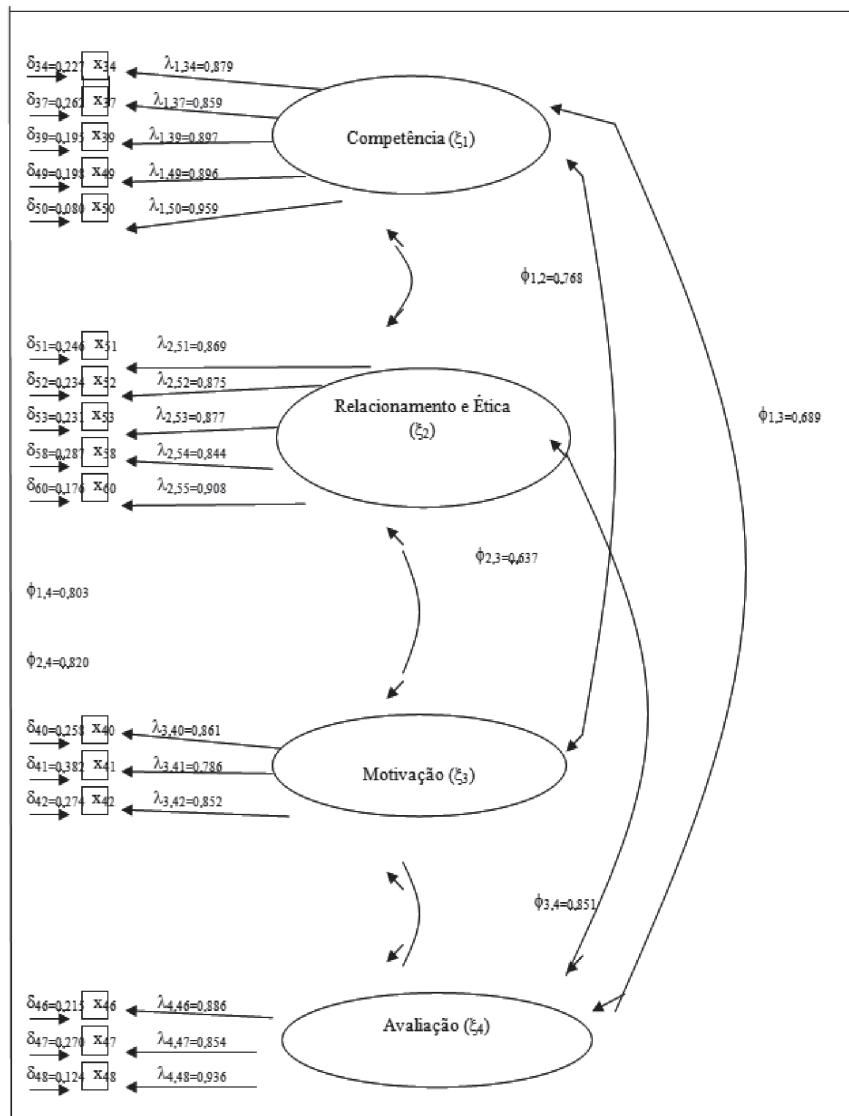
Fonte: Elaborada pelos autores (2005).

Notas: Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade composta do fator. (2) * Todos os lâmbdas (λ) são estatisticamente significativos ao nível de 1%.

(3) ** Variância extraída da variável latente.

VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

Diagrama 2 - Diagrama de caminhos das características da qualidade docente



Fonte: Elaborada pelos autores (2005).

Quanto à significância dos indicadores, todos são positiva e significativamente relacionados com seus fatores latentes específicos. Esse fato pode ser constatado dividindo cada lâmbda estimado pelo seu respectivo erro padrão. Todos os *t student*, são estatisticamente significativos em níveis menores que 1%. Portanto, todos os indicadores estão significativamente relacionados com seus fatores latentes específicos, comprovando a relação proposta entre os indicadores e os fatores latentes.

No que tange à confiabilidade dos indicadores, todos apresentam erros de medida (*theta delta*) menores que 0,5. Todos os indicadores exibem confiabilidade bastante elevada, indicando que eles explicam uma alta porcentagem da variância das variáveis latentes. A confiabilidade composta de cada fator latente é elevada, superando em todos os casos o limite mínimo de aceitação de 0,7. Desse modo, pode-se afirmar que os indicadores medem consistentemente seus respectivos construtos. Em todos os fatores, a variância extraída supera a cota mínima que é de 0,5, indicando que a variância dos indicadores associada a esse fator latente é suficientemente explicada por ele.

Todas as relações causais, apresentadas no diagrama 2, são positivas, significativas e de diferente intensidade. O indicador x_{50} , (conduz o ensino da disciplina com qualidade), é o indicador mais influenciado pelo fator “competência”, comprovou-se isso pela elevada saturação desse indicador no fator. O indicador x_{60} , (valoriza a cooperação e a solidariedade em sala de aula) é o indicador que mais fortemente está saturado com o fator “relacionamento e ética”. O indicador x_{40} , (estimula a utilização de bibliografia atualizada) é o indicador mais fortemente saturado no fator “motivação”, enquanto o indicador x_{48} , (planeja avaliações compatíveis com os objetivos e conteúdos ministrados), é o mais altamente explicado pelo fator “avaliação”. Todos os indicadores medidos evidenciam baixos erros de medida. Logo, comprovam a alta confiabilidade evidenciada. As correlações entre os fatores latentes confirmam as obtidas na análise fatorial exploratória. As correlações entre os fatores são positivas e significativas e evidenciam que o sucesso do professor depende de suas habilidades nesses fatores.

Em síntese, há evidências de que o modelo de medida com quatro fatores correlacionados ajusta-se satisfatoriamente aos dados, possibilitando considerar os indicadores como válidos e confiáveis.



6 CONCLUSÃO

O objetivo desta pesquisa foi testar as propriedades de confiabilidade e validade do instrumento de avaliação da docência da Universidade do Sul de Santa Catarina. Inicialmente estimou-se o modelo de medida com oito fatores correlacionados. O valor do χ^2 desse modelo é 3474,20. A análise detalhada do modelo conduziu para a sua reavaliação. Os estudos exploratórios evidenciaram, após a remoção de alguns indicadores, os quais traziam mais ruídos do que informações, que a estrutura da escala atual pode ser explicada com pouca perda de informação, por apenas quatro fatores latentes correlacionados. Essa hipótese foi submetida à prova por meio dos modelos causais e resultou um χ^2 de 375,235. A proporção no aumento do ajuste, quando se compara o χ^2 do modelo com oito fatores correlacionados com o χ^2 do modelo com quatro fatores correlacionados, é aproximadamente 0,9. Os índices de qualidade de ajuste global, incremental e os de parcimônia, obtidos pelo modelo com quatro fatores latentes correlacionados, são similares aos modelos de avaliação da docência de Ramos (1997) e do modelo de avaliação de curso de Capellera e Veciana (2000).



Em síntese, os diversos índices de confiabilidade dos indicadores e dos construtos decorrentes dos estudos confirmatórios foram superiores a 0,7, indicando consistência nas avaliações dos discentes da UNISUL. Quanto aos indicativos de validade de construto, os índices incrementais de ajuste TLI e BBI foram superiores a 0,9 em todos os estudos confirmatórios, indicando que os alunos reconhecem que o instrumento em análise apresenta quatro dimensões correlacionadas. As dimensões validadas da escala são: competência, com cinco indicadores; relacionamento e ética, com cinco indicadores; motivação, com três indicadores e, por último, a dimensão avaliação, com três indicadores.



Entende-se que a qualidade do processo de ensino e aprendizagem é bastante complexa e, por isso, não pode ser resumido apenas aos quatro fatores validados nesta pesquisa. Por esse motivo, faz-se necessária uma adequação do instrumento à luz das teorias do processo de ensino e de aprendizagem.

A metodologia, aqui empregada para estimar o grau de confiabilidade e de validade de instrumento de avaliação da docência pelos discentes, pode ser utilizada em qualquer ramo de atividade que utilize escala de multiitens, desde que a escala seja valorada por uma amostra ampla de respondentes.

Do ponto de vista das limitações desta pesquisa, é preciso salientar que, mesmo tendo assegurado que o instrumento de avaliação apresente índices de

confiabilidade e validade razoável para um conjunto de discentes num determinado corte do tempo, não assegura garantias plenas de se estar medindo o nível de satisfação dos alunos com um instrumento tecnicamente correto. O processo de validação de um questionário implica um conjunto de decisões que se apóia em contrastes de hipóteses corretamente formuladas e testadas a partir dos dados provenientes de processos avaliativos distintos e em contextos diversos. O acúmulo de todos os indicadores de confiabilidade e de validade, oriundos da aplicação do instrumento em diferentes contextos educacionais, fornecerá o suporte empírico da sua adequação.

Por fim, esta pesquisa sugere que sejam realizadas pesquisas sobre o efeito das variáveis “tamanho das classes”, “dificuldade da disciplina”, “alunos poucos preparados para prosseguirem no curso elegido”, “idade dos alunos”, “sexo dos alunos”, “tipo de disciplinas”, “notas”, “semestre letivo” e outras variáveis que possam provocar vieses negativos nos níveis de satisfação dos alunos. Recomenda-se, também, que sejam realizadas pesquisas junto aos professores e alunos para investigar o efeito dos resultados das avaliações dos professores no melhoramento da qualidade do seu ensino, ou seja, investigar a validez de uso dos resultados das avaliações dos alunos, por parte dos docentes. E finalmente, recomenda-se que sejam considerados outros fatores institucionais que podem interferir no desempenho, tais como infra-estrutura de laboratórios, recursos didáticos, bibliografia disponível, entre outros.

REFERÊNCIAS

- AKAIKE, H. Factor analysis and AIC. *Psychometrika*, California, n. 52, p. 317-332, 1987.
- ALMEIDA, Alzira Jerônimo de Melo. O professor e a valorização de sua atividade docente. In: FELTRAN, Regina Célia de Santis (Org.). **Avaliação na educação superior**. Campinas: Papirus, 2002, p. 19-190.
- AUBRECHT, J. D. **Are student ratings of teacher effectiveness valid?** (IDEA paper n. 2). Manhattan: Kansas State University, Center for Faculty Evaluation and Development, 1979. (ERIC document reproduction service n. ED 202410).

VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

BENTLER, P. M.; BONETT, D. G. Significance tests and goodness of fit in the analysis of covariance structures. **Psychological Bulletin**, New York, v. 88, n. 3, p. 588-606, 1980.

BOLLEN, K. A. **Structural equations with latent variables**. New York: Wiley, 1989.

BROWNE, M. W; CUDEC, R. Alternative ways of assessing model fit. In: BOLLEN, K. A., LONG, J. S. (Orgs.). **Testing structural equation models**. Newbury Park, CA: Sage, 1993. p. 136-162.

CAPELLERAS, Joan-Lluíz; VECIANA, J. M^a. Calidad de servicio en la enseñanza universitaria: desarrollo y validación de una escala de medida. In: CONGRESSO NACIONAL DE ACEDE, 11, 2000. Disponível em: <www.fnac.es>. Acesso em: 18 set. 2003.

COHEN, P. A. Student rating of instruction and student achievement: a metaanalysis of multisector validity studies. **Review of Educational Research**, Chicago, v. 3, p. 281-309, 1981.

CRONBACH, L. J. Coefficient alfa and the internal structure of tests. **Psychometrika**, California, v. 16, p. 297-334, 1951.

GARCIA, S. D. B.; MARTINEZ, T. L. Análisis de ecuaciones estructurales. In: GREIMEL-FUHRMANN, B.; GEYGER, A. Student's evaluation of teachers and instructional quality analysis of relevant factors based on empirical evaluation research. **Assessment Evaluation in Higher Education**, London, v. 28, n. 3, p. 229-238, 2003.

GREIMEL-FUHRMANN, B.; GEYGER, A. Student's evaluation of teachers and instructional quality analysis of relevant factors based on empirical evaluation research. **Assessment Evaluation in Higher Education**, London, v. 28, n. 3, p. 229-238, 2003.

HAIR, J.F. et al. **Análises multivariantes**. 5. ed. Madrid: Prentice Hall Iberia, 1999.

HAIR JR.; J. F. et al. **Fundamentos de métodos de pesquisa em administração**. Porto Alegre: Bookman, 2005.

JÖRESKOG, K. G. A general approach to confirmatory maximum likelihood factor analysis. **Psychometrika**, California, n. 34, p. 183-202, 1969.

LOESCH, C.; HOELTGEBAUM, M. **Métodos estatísticos multivariados aplicados à economia de empresas.** Blumenau: Nova Letra, 2005.

MALHOTRA, N. K. **Pesquisa de marketing:** uma orientação aplicada. 3. ed. Porto Alegre: Bookman, 2001.

MARSH, H. W. et al. Goodness-of-fit indexes in confirmatory factor analysis: the effect of sample size. **Psychological Bulletin**, New York, v. 103, n. 3, p. 391-410, 1988.

MARSH, H. W; BALLA, J. R.; MCDONALD, R. P. Goodness-of-fit indexes in confirmatory factor analysis: the effect of sample size. **Psychological Bulletin**, New York, v. 103, n. 3, p. 391-410, 1988.

MATEO G, M. A.; FERNÁNDEZ, S. J. Análisis confirmatório de la estructura dimensional de um cuestionário para la evaluación de la calidad de la enseñanza. **Investigaciones Psicológicas**, Madrid, n. 11, p. 73-80, 1992.



MARSH, H. W. Student's evaluation of teaching effectiveness: the stability of mean ratings of the same teacher over a 13-year period. **Teaching & Teacher Education**, n. 7, p. 303-314, 1991.



MARDIA, K. A. Measures of multivariate shewness and kurtosis with applications. **Biometrika**, Oxford, 105, p. 430-445, 1989.

MCDONALD, R. P.; MARSH, H. W. Choosing a multivariate model: noncentrality and goodness of fit. **Psychological Bulletin**, New York, n. 107, p. 247-55, 1990.

MOREIRA, Daniel A. **Avaliação do professor universitário pelo aluno:** possibilidades e limitações. (Tese de doutorado) - Universidade de São Paulo, 1986.

MULAIK, S. A. et al. Evaluation of goodness-of-fit indexes for structural equation models. **Psychological Bulletin**, New York, v. 105, n. 3, p. 340-445, 1989.

NUNNALLY, J. C. **Psychometric theory.** New York: McGraw-Hill, 1978.

PEDHAZUR, E. J ; SCHMELKIN L. P. **Measurement, design, and analysis.** An integrated approach. Hillsdale, N. J.: Lawrence Erlbaum, 1991.



VALIDADE E CONFIABILIDADE DE INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA SOB A ÓTICA DOS
MODELOS DE EQUAÇÃO ESTRUTURAL

RAMOS, J. M. G. El análisis factorial confirmatorio aplicado a la investigación pedagógica no-experimental. **Bordón**, Madrid, n. 267, p. 245-267, 1987.

_____. Validacion de construto en el ambito pedagogico. **Revista Española de Pedagogia**, Madrid, v. 174, p. 535-554, 1986.

_____. Análises factorial confirmatória en la validación del constructo competencia docente del professor universitario. **Bordón**, Madrid, v. 49, n. 4, p. 361-391, 1997.

TOURÓN, J. La validacion de construto: sua aplicacion al CEED (Cuestionário para la evaluacion de la eficacia docente. **Bordón**, Madrid, p. 735-756, 1989.

TUCKER, L. R; LEWIS, C. A reliability coefficient for maximum likelihood factor analysis. **Psychometrika, California**, v. 38, n. 1, p. 1-10, 1973.



VAL, J. C. Análisis factorial confirmatorio de las características de calidad docente universitaria (solução LISREL). **Bordón**, Madrid, v. 4, n. 46, p. 389-405, 1994.



ANEXO - ITENS DO QUESTIONÁRIO

01. Considera que seu curso tem qualidade
02. Tem informações regulares sobre o seu curso.
03. Conhece o projeto do seu curso como um todo.
04. O seu curso está exigindo o suficiente para a sua formação profissional.
05. Considera sua universidade de qualidade.
06. Você tem orgulho de estudar em sua universidade.
07. Seu curso tem uma boa imagem na sociedade.
08. O formado em seu curso tem boa aceitação no mercado de trabalho.
09. Seu curso goza de prestígio junto aos estudantes de sua universidade.
10. Encontram na biblioteca as referências que procuram.
11. A biblioteca oferece atendimento de qualidade.
12. Está satisfeito com o acervo da universidade
13. A secretaria geral e protocolo demonstram competência nas suas funções.
14. O serviço de Xerox e mecanografia correspondem as suas necessidades.
15. O setor de apoio ao estudante é competente no que diz respeito a bolsas.
16. Você recebe informação sobre o setor de apoio ao estudante.
17. Você conhece o setor de apoio aos estágios da sua universidade.
18. Você dispõe de internet em casa.
19. Você está satisfeito com o serviço de internet de sua universidade.
20. O ambiente do campus é propício ao desenvolvimento de atividades acadêmicas.
21. As salas de aula são adequadas para o desenvolvimento das aulas.
22. As condições dos banheiros são adequadas.
23. Você considera que existe segurança no campus.
24. Você está satisfeito com os serviços de lanchonete.
25. Oferece horário de atendimento acessível aos alunos
26. Comunica-se com a turma.
27. Revela ser um líder democrático.
28. Demonstra competência na resolução dos problemas.
29. Promove atividades extracurriculares.
30. Informa sobre eventos do curso.
31. Indicaria o seu coordenador como modelo de atuação.

32. É pontual, cumpre os horários das aulas
33. É assíduo, comparece às aulas.
34. Administra o tempo em sala de forma produtiva.
35. Apresentou o plano de ensino.
36. Está desenvolvendo o plano de ensino.
37. Revela domínio de conteúdo na sua disciplina.
38. Expressa o conteúdo das aulas com linguagem acessível aos alunos.
39. A bagagem intelectual é estimulante para os alunos.
40. Estimula a utilização de bibliografia atualizada.
41. Estimula a produção científica promovendo atividades de pesquisa na disciplina.
42. Incentiva a leitura de livros, texto, jornais e revistas complementares à aula.
43. Adota procedimentos facilitadores da aprendizagem.
44. Oportuniza reflexões críticas sobre os conteúdos.
45. Dá atenção às dúvidas apresentadas pelos alunos.
46. Redige provas ou verificações de aprendizagem de forma clara.
47. Informa sobre os critérios adotados na avaliação.
48. Planeja avaliações compatíveis com os objetivos e conteúdos ministrados.
49. Trabalha a disciplina com nível de exigência suficiente.
50. Conduz o ensino da disciplina com qualidade.
51. Tem boa relação com os alunos.
52. Valoriza a participação do aluno nas aulas.
53. Estimula o bom relacionamento da classe.
54. Demonstra conhecer o projeto do curso.
55. Situa o aluno sobre a importância da sua disciplina para o curso.
56. Demonstra interesse e compromisso com a Unisul.
57. Demonstra conhecimento da Unisul no seu campus.
58. Respeita o aluno como pessoa.
59. Demonstra coerência entre o discurso e sua prática como professor.
60. Valoriza a cooperação e a solidariedade em sala de aula.
61. Indicaria esse professor a outras turmas.
62. Você se sente tranquilo para opinar sobre esse professor.

