

Universidade Federal de Santa Catarina
Programa de Pós-Graduação em
Engenharia de Produção

CONFIABILIDADE E VALIDADE ESTATÍSTICAS DA AVALIAÇÃO
DOCENTE PELO DISCENTE:
PROPOSTA METODOLÓGICA E ESTUDO DE CASO

Tese

Amilton Barreto de Bem

Florianópolis

2004

Amilton Barreto de Bem

**CONFIABILIDADE E VALIDADE ESTATÍSTICAS DA AVALIAÇÃO
DOCENTE PELO DISCENTE:
PROPOSTA METODOLÓGICA E ESTUDO DE CASO**

Tese apresentada ao Programa de Pós-Graduação
em Engenharia de Produção da Universidade
Federal de Santa Catarina para a obtenção do
título de Doutor em Engenharia de Produção.

Orientador: Prof. Edgar Augusto Lanzer, Ph.D.

Florianópolis

2004

Amilton Barreto de Bem

**CONFIABILIDADE E VALIDADE ESTATÍSTICAS DA AVALIAÇÃO
DOCENTE PELO DISCENTE:
PROPOSTA METODOLÓGICA E ESTUDO DE CASO**

Esta tese foi julgada e aprovada para a obtenção do grau de Doutor em Engenharia de Produção no Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina.

Florianópolis, novembro de 2004.

Prof. Nelson Pacheco Paladini
Coordenador do Programa

BANCA EXAMINADORA

Prof. Orientador Edgar Augusto Lanzer, Ph.D.

Prof. Dra. Ana Lúcia Miranda Lopes

Prof. Dr. Eduardo Búrigo de Carvalho

Prof. Dr. Jailson Coelho

Prof. Dr. Pedro Alberto Barbetta

Prof. Dr. Elmo Tambosi Filho

Prof. Dr. Wesley Vieira da Silva

Agradecimentos

À Universidade Federal de Santa Catarina, através do Programa de Pós-Graduação em Engenharia de Produção e Sistemas.

À Universidade do Sul de Santa Catarina, pelo apoio e incentivo à capacitação profissional dos seus professores.

Um agradecimento especial à equipe de Avaliação Institucional e Clima Organizacional da UNISUL, que me oportunizou a elaboração deste projeto.

Aos amigos Eduardo Búrigo de Carvalho, Jailson Coelho, Elmo Tambosi Filho que contribuíram, participando das discussões.

À minha querida esposa e filhos que compreenderam o meu afastamento para dedicação a este estudo.

Por fim, à meu orientador professor Edgar Augusto Lanzer pela compreensão e oportunidade de realização.

Resumo

DE BEM, Amilton Barreto. **Confiabilidade e validade estatísticas da avaliação docente pelo discente**: proposta metodológica e estudo de caso. 2004. 296f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, Florianópolis.

Este trabalho propõe uma metodologia para testar as propriedades de confiabilidade e de validade de instrumento de avaliação da docência pelos discentes em instituições de ensino. Nesse se discutiu o uso das técnicas analíticas de análise fatorial exploratória como ferramenta de análise de dados para auxiliar os especialistas, em medição educacional, no estudo da confiabilidade e da validade da dimensionalidade dos instrumentos de avaliação da docência. Para testar as hipóteses decorrentes dos estudos exploratórios é apresentada a metodologia de análise fatorial confirmatória, a qual utiliza-se de uma série de índices de ajustes absolutos, incrementais e de parcimônia para validar as hipóteses decorrentes dos estudos exploratórios. A metodologia foi empregada e julgada adequada para investigar as propriedades estatísticas de confiabilidade e validade das oito dimensões do instrumento de avaliação da docência da Universidade do Sul de Santa Catarina - UNISUL.

Palavras-chave: instrumento, avaliação, confiabilidade, validade.

Abstract

DE BEM, Amilton Barreto. **Confiabilidade e validade estatísticas da avaliação docente pelo discente**: proposta metodológica e estudo de caso. 2004. 296f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, Florianópolis.

This work proposes a methodology in order to test the reliability and validity properties of the faculty's evaluation tools by the student body on teaching institutions. The analytical techniques of exploratory factorial analysis were discussed, as a data analysis tool to help specialists measure the education regarding the reliability and validity study of the dimensionality of the faculty's evaluation tools. In order to test the ongoing hypothesis of the exploratory studies, the confirmatory factorial analysis methodology is presented, and it uses a number of absolute adjustment indexes, incremental and parsimonious ones so as to validate the ongoing hypothesis of the exploratory studies. The methodology was used and judged to be adequate to investigate the statistical properties of reliability and validity of eight evaluation tool dimensions of UNISUL's (Universidade do Sul de Santa Catarina) faculty.

Key Words: Tool, evaluation, reliability, validity.

Sumário

LISTA DE GRÁFICOS	14
LISTA DE QUADROS	15
LISTA DE DIAGRAMAS	16
LISTA DE FLUXOGRAMAS	17
LISTA DE TABELAS.....	18
1 INTRODUÇÃO.....	21
1.1 Justificativa e importância do trabalho.....	25
1.2 Problema e hipóteses básicas da pesquisa	26
1.2.1 Problema	26
1.2.2 As hipóteses básicas	28
1.3 Objetivos.....	29
1.3.1 Geral	29
1.3.2 Específicos	30
1.4 Estrutura do trabalho	30
2 MULTIDIMENSIONALIDADE DOS INSTRUMENTOS DE AVALIAÇÃO DA DOCÊNCIA.....	33
2.1 Introdução	33
2.2 Utilidade das avaliações dos alunos	35
2.3 Dimensionalidade, confiabilidade e validade	38
2.4 Paradigmas norteadores da avaliação da docência.....	39
2.5 Teorias implícitas dos alunos	41

2.6 Dimensionalidade dos instrumentos de avaliação da docência.....	43
3 ASPECTOS TEÓRICOS DA CONFIABILIDADE DAS AVALIAÇÕES DOS DOCENTES PELOS DISCENTES	59
3.1 Introdução	59
3.2 Erros de medição.....	60
3.3 Fontes perturbadoras da confiabilidade	61
3.4 O modelo linear de pontuação	63
3.5 Medição da confiabilidade	65
3.6 Medição da confiabilidade – teoria da generalização	68
3.7 Medição da confiabilidade – o coeficiente α de Cronbach.....	70
3.8 Relação entre validade e confiabilidade.....	72
3.9 Pesquisas sobre a confiabilidade da avaliação dos docentes pelos discentes.....	74
3.10 Estabilidade das avaliações	79
3.11 Generalização das avaliações	81
4 VALIDADE DAS AVALIAÇÕES DOS DOCENTES PELOS DISCENTES	84
4.1 Introdução	84
4.2 Tipos de validade	86
4.3 Validação de construto.....	91
4.4 Validade de construto – enfoque de Ramos.....	94
4.4.1 Explicitação da teoria sobre o construto.....	95
4.4.2 Formulação das hipóteses derivadas da rede teórica e dos estudos exploratórios sobre o construto	96
4.4.3 Submetendo as hipóteses à prova através de estudos confirmatórios	97

4.4.4 Avaliação e integração dos resultados.....	97
4.4.5 Formulação de novas hipóteses.....	98
4.5 Pesquisas sobre a validade da avaliação dos docentes pelos discentes	100
4.5.1 Validade de construto – experimento de multiseção.....	102
4.5.2 Validade de construto – modelos de equação estrutural.....	107
4.5.3 Validade de construto – matrix de multitração-multimétodo	113
4.6. Outros critérios relacionados com a avaliação dos estudantes.....	115
4.7 Fontes possíveis de viés nas avaliações dos alunos	120
5 AVALIAÇÃO DAS PROPRIEDADES DO INSTRUMENTO DE AVALIAÇÃO	
MEDIANTE MODELOS DE ANÁLISE FATORIAL EXPLORATÓRIO	127
5.1 Introdução	127
5.2 O processo de análise fatorial	133
5.2.1 Problema de investigação	134
5.2.2 Seleção do tipo de análise fatorial.....	134
5.2.3 Projeto de investigação.....	135
5.2.4 Pressupostos para a adequação do modelo	136
5.2.5 Seleção de um modelo de fatores.....	139
5.2.6 Especificação da matriz fatorial.....	139
5.2.7 Seleção de um método de rotação.....	140
5.2.8 Interpretação da matriz fatorial de rotação	141
5.2.9 Reespecificação do modelo fatorial.....	141
5.2.10 Validação da matriz fatorial.....	142
5.2.11 Análise fatorial exploratória <i>versus</i> confirmatória.....	142

5.2.12 O uso da técnica de análise fatorial exploratório para a validade de construto	143
---	-----

6 AVALIAÇÃO DAS PROPRIEDADES DO INSTRUMENTO DE AVALIAÇÃO

MEDIANTE MODELOS DE ANÁLISE FATORIAL CONFIRMATÓRIO.....	146
--	------------

6.1 Introdução	146
-----------------------------	------------

6.2 Fases para o desenvolvimento de um modelo de equações estruturais	148
--	------------

6.3 Especificação do modelo.....	150
---	------------

6.3.1 Desenvolvimento de um modelo baseado na teoria	150
--	-----

6.3.2 A construção de um diagrama de relações.....	150
--	-----

6.3.2.1 Modelo estrutural	153
---------------------------------	-----

6.3.3 Modelo de medida de indicadores exógenos e endógenos	155
--	-----

6.4 Identificação do modelo de medida	159
--	------------

6.5 Estimação do modelo de medida.....	160
---	------------

6.5.1 Eleição do tipo de matriz	161
---------------------------------------	-----

6.5.2 Estimação dos parâmetros do modelo	163
--	-----

6.5.2.1 Método de estimação.....	164
----------------------------------	-----

6.5.2.2 Processos de estimação.....	165
-------------------------------------	-----

6.6 Avaliação e interpretação do modelo de medida.....	166
---	------------

6.6.1 Medidas de ajuste absoluto	168
--	-----

6.6.1.1 Texto do χ^2	169
---------------------------------	-----

6.6.1.2 Parâmetros de não centralidade.....	169
---	-----

6.6.1.3 Índice de bondade de ajuste (GFI)	170
---	-----

6.6.1.4 Resíduo quadrático médio (RMSR)	171
---	-----

6.6.2 Medidas incrementais de ajuste	171
--	-----

6.6.2.1 Índice de bondade de ajuste ajustado (AGFI).....	172
--	-----

6.6.2.2 Índice de Tucker-Lewis (TLI).....	172
6.6.2.3 Índice de ajuste normal (BBI).....	173
6.6.3 Medidas de ajuste de parcimônia	173
6.6.3.1 Índice de ajuste normalizado de parcimônia (PNFI).....	174
6.6.3.2 Índice de bondade de ajuste de parcimônia (PGFI)	174
6.6.3.3 Qui-quadrado normalizado	175
6.6.3.4 Critério de informação de Akaike (AIC)	175
6.6.4 Ajuste do modelo de medida.....	178
6.6.5 Confiabilidade composta do construto	178
6.6.6 Variância extraída.....	179
6.7 Interpretação e modificação do modelo.....	180

7 APLICAÇÃO DA METODOLOGIA AO INSTRUMENTO DE AVALIAÇÃO

DA DOCÊNCIA DA UNISUL	184
7.1 Introdução	184
7.2 A entidade.....	184
7.3 Tipo de pesquisa e base de dados.....	186
7.4 Objeto da pesquisa	188
7.5 Método da pesquisa	188
7.6 Variáveis do estudo	189
7.7 Tratamento dos dados.....	191
7.8 Hipóteses da pesquisa.....	192
7.9 Limitações da pesquisa.....	201
7.10 Técnicas de análises de dados e software para análises	202

8 ANÁLISE DAS PROPRIEDADES ESTATÍSTICA DE CONFIABILIDADE E DE VALIDADE DO INSTRUMENTO DA UNISUL “AVALIAÇÃO PELOS ALUNOS DO DESEMPENHO NA DISCIPLINA”	203
8.1 Introdução	203
8.2 Tratamento dos dados	204
8.3 Estimativas da confiabilidade dos indicadores e das dimensões do instrumento de avaliação de disciplinas da UNISUL	206
8.4 Identificação do modelo com oito fatores não correlacionados e do modelo com oito fatores correlacionados e tamanhos de amostras	208
8.5 Estimação do modelo com oito fatores latentes não correlacionados e do modelo com oito fatores latentes correlacionados	210
8.6 Modificação do modelo	220
8.7 Estimação do modelo com um fator latente e do modelo com quatro fatores latentes correlacionados	233
8.8 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – geral UNISUL	236
8.9 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – <i>campus</i> Tubarão	241
8.10 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – <i>campus</i> Araranguá	246
8.11 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – <i>campus</i> Pedra Branca	250
8.12 Comparação dos índices de adequação do ajuste do modelo: geral UNISUL e de outros pesquisadores	256
8.13 Validez de critério da escala	257

8.14 Validade da escala para diagnosticar tipologias de docentes.....	263
9 CONCLUSÕES.....	266
REFERÊNCIAS.....	272
ANEXOS	281
ANEXO I: Questionário Discente	282
ANEXO II: Matriz de correlação geral Unisul (estudos exploratórios)	283
ANEXO III: Matrix de correlação <i>campus</i> Tubarão (estudos exploratórios)	284
ANEXO IV: Matrix de correlação <i>campus</i> Araranguá (estudos exploratórios).....	285
ANEXO V: Matrix de correlação <i>campus</i> Pedra Branca (estudos exploratórios).....	286
ANEXO VI: Matrix de correlação (análise fatorial confirmatório)–Unisul Geral.....	287
ANEXO VII: Matrix de correlação (análise fatorial confirmatório)– <i>Campus</i> Tubarão	288
ANEXO VIII: Matrix de correlação (análise fatorial confirmatório)– <i>Campus</i> Araranguá ...	289
ANEXO IX: Matrix de correlação (análise fatorial confirmatório)– <i>Campus</i> Pedra Branca .	290
ANEXO X: Matrix de correlação (análise confirmatório - modificado)–Unisul Geral.....	291
ANEXO XI: Matrix de correlação (análise confirmatório - modificado)– <i>Campus</i> Tubarão	292
ANEXO XII: Matrix de correlação (análise confirmatório - modificado)– <i>Campus</i> Pedra Branca.....	293
ANEXO XIII: Matrix de correlação (análise confirmatório - modificado)–Geral Unisul.....	294
ANEXO XIV: Matrix de correlação (análise confirmatório - modificado)– <i>Campus</i> Tubarão	295
ANEXO XV: Matrix de correlação (análise confirmatório - modificado)– <i>Campus</i> Pedra Branca.....	296

Lista de Gráficos

Gráfico I: Distribuições dos indicadores avaliados. Geral Unisul.....	204
Gráfico II: Distribuições dos indicadores avaliados. Geral Unisul (continuação)	205
Gráfico III: Gráfico de probabilidade normal dos resíduos do modelo com oito fatores correlacionados, geral Unisul	214
Gráfico IV: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados – geral Unisul.....	237
Gráfico V: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados – campus Tubarão	242
Gráfico VI: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados – campus Araranguá.....	247
Gráfico VII: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados – campus Pedra Branca	252
Gráfico VIII: Médias dos grupos – geral Unisul	263

Lista de Quadros

Quadro 1: Comparação entre confiabilidade e validade, segundo o tipo de erro.....	73
Quadro 2: Tipos de validade requerida para com os instrumentos de avaliação da docência.	91
Quadro 3: Quadro comparativo entre Análise Fatorial Exploratória x Confirmatória	143
Quadro 4:Resumo das medidas de adequação do ajuste.....	177
Quadro 5: Dimensões do bloco “Desempenho do Docente”	190

Lista de Diagramas

Diagrama 1: Representação do modelo fatorial com K indicadores e dois fatores comuns .	128
Diagrama 2: Diagrama de caminhos das características da qualidade docente(três fatores latentes correlacionados)	152
Diagrama 3: Diagrama de caminhos das características da qualidade docente(oito fatores latentes não correlacionados).....	193
Diagrama 4: Diagrama de caminhos das características da qualidade docente(oito fatores latentes correlacionados)	196
Diagrama 5: Diagrama de caminhos das características da qualidade docente – um fator latente.....	230
Diagrama 6: Diagrama de caminhos das características da qualidade docente com quatro fatores latentes correlacionados	232
Diagrama 7: Diagrama de caminhos das características da qualidade docente – Geral da UNISUL	262

Lista de Fluxogramas

Fluxograma 1: Aspectos importantes da validade de construto	94
Fluxograma 2: Passos do processo de validação de construto	99
Fluxograma 3: Passos utilizados na extração dos fatores do modelo de análise fatorial exploratória	133
Fluxograma 4: Passos utilizados na confirmação dos fatores decorrentes das análises exploratórias.....	149

Lista de Tabelas

Tabela 1: Universidade do estado de Kansas Indicadores de Confiabilidade do Instrumento IDEA.....	75
Tabela 2: Universidade de Sydney – Austrália Indicadores da confiabilidade do Instrumento SEEQ.....	75
Tabela 3: Instrumentos de Avaliação da docência universitária Indicadores da confiabilidade	76
Tabela 4: Confiabilidade dos fatores dos instrumentos SEEQ e ENDEAVOR – Universidades: Austrália e Espanha.....	77
Tabela 5: Confiabilidade dos fatores do CEQ Universidade de Melbourne	78
Tabela 6: Confiabilidade das dimensões da escala	79
Tabela 7: Correlações do Estudo de Hogan (1973), Bausell <i>et al.</i> (1975) e Marsh (1982)	82
Tabela 8: Coeficientes de correlação entre as dimensões do instrumento e qualidade global e satisfação	110
Tabela 9: Coeficientes de convergência da matriz de multitraço-multimétodo dos instrumentos de avaliação de Marsh (1982) e de Frey <i>et al.</i> (1975).....	114
Tabela 10: Coeficiente de validade convergente das dimensões do instrumento SEEQ, evidenciados a partir das respostas dos alunos e dos professores	117
Tabela 11: Número de cursos e de alunos por nível de ensino – 2002	186
Tabela 12: Matriculados na UNISUL, no segundo semestre letivo de 2002	187
Tabela 13: Confiabilidade dos indicadores e das dimensões do instrumento de avaliação de disciplinas da UNISUL – Geral	207

Tabela 14: Confiabilidades das dimensões do instrumento de avaliação de disciplinas da UNISUL, estratificadas segundo o <i>campus</i>	208
Tabela 15: Graus de liberdade dos modelos de medida apresentado nos diagramas 3 e 4....	209
Tabela 16: Números de disciplinas utilizadas para comprovação dos modelos, expostos nos fluxogramas 5 e 6.....	210
Tabela 17: Índices de adequação do ajuste dos modelos com oito fatores não correlacionados e oito fatores correlacionados	211
Tabela 18: Valores dos coeficientes lâmbda (λ), erro padrão e erro de medida dos indicadores - Geral da Unisul.....	213
Tabela 19: Índices de adequação de ajuste dos modelos com oito dimensões ou fatores não correlacionados e oito dimensões ou fatores correlacionados, estratificados segundo o <i>campus</i>	216
Tabela 20: Correlações entre os fatores latentes.....	218
Tabela 21: Confiabilidade composta e variância extraída dos fatores latentes correlacionados – geral UNISUL	219
Tabela 22: Medidas do grau de associação entre os indicadores.....	221
Tabela 23: Matriz fatorial rotacionada – geral UNISUL	222
Tabela 24: Matriz fatorial rotacionada – <i>Campus</i> Tubarão.....	223
Tabela 25: Matriz fatorial rotacionada – <i>Campus</i> Araranguá	224
Tabela 26: Matriz fatorial rotacionada – <i>Campus</i> Pedra Branca.....	225
Tabela 27: Confiabilidade dos indicadores decorrentes dos estudos exploratórios. Geral da UNISUL e estratificado segundo o <i>Campus</i>	226
Tabela 28: Tamanho das amostras utilizadas para comprovação dos modelos apresentados nos diagramas 5 e 6.....	234
Tabela 29: Graus de liberdade dos modelos de medida apresentados nos diagramas 5 e 6...	235

Tabela 30: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores	236
Tabela 31: Análise fatorial confirmatório. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – Geral UNISUL	240
Tabela 32: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores	241
Tabela 33: Análise fatorial confirmatório. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – <i>Campus Tubarão</i>	245
Tabela 34: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores	246
Tabela 35: Análise fatorial confirmatório. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – <i>Campus de Araranguá</i>	250
Tabela 36: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores correlacionados. Estimação por máxima verossimilhança – <i>Campus: Pedra Branca</i>	251
Tabela 37: Análise fatorial confirmatória. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do fator. Estimação por máxima verossimilhança – Modelo com quatro fatores correlacionados: <i>Campus: Pedra Branca</i>	255
Tabela 38: Comparação dos índices de adequação do ajuste	256
Tabela 39: Análise de regressão	258
Tabela 40: Correlações entre as variáveis critérios e os indicadores das dimensões validadas da escala	260
Tabela 41: Valores médios por cluster e análise de variância	265

1 INTRODUÇÃO

Nos anos 80, a realização de seminários como o encontro nacional para tratar sobre “Avaliação do Ensino Superior” (BRASIL, 1988) foi o início de uma nova era no âmbito da educação superior no Brasil. Mas, somente na década de 90 foram estruturadas e implementadas atividades de avaliação, tais como o “Programa de Avaliação Institucional das Universidades Brasileiras” (PAIUB), o “Exame Nacional dos Cursos de Graduação” (ENC) e as “Comissões de Especialistas” (Ces). Todos esses programas indicam que a avaliação da qualidade dos cursos de graduação ocupa relevante espaço no Brasil.

Paralelo aos programas desencadeados pelo MEC foi lançada e aderido por algumas universidades brasileiras a “Proposta de Modelo de Avaliação Institucional para as Universidades Brasileiras”, apresentada pelo Conselho de Reitores das Universidades Brasileiras (CRUB). Tal conselho partiu da hipótese de que o êxito de uma instituição de ensino superior está relacionado com a sua capacidade de autopercepção; ou seja, o sucesso de uma instituição de ensino está diretamente associado com a sua capacidade de se auto-avaliar (CONSELHO DOS REITORES, 2000).

Na Universidade do Sul de Santa Catarina (UNISUL), a avaliação institucional surgiu em 1995 com a designação de uma comissão de Avaliação Institucional, institucionalizada pelo Reitor, com o objetivo de disseminar, no **Campus**, a cultura da avaliação, desenvolver instrumentos fidedignos e válidos para a obtenção de informações úteis que possibilitassem o desenvolvimento de ações institucionais, visando à garantia da qualidade do ensino de graduação, dos serviços de apoio ao ensino e de gestão.

Por considerar a avaliação um processo de forte conteúdo ético, a Comissão de Avaliação Institucional da UNISUL, orientada por princípios da **legitimidade** - a avaliação é realizada em benefício do avaliado; **da utilidade** - a avaliação deve corroborar com a

melhoria do desempenho do docente e do curso; **da factibilidade** - a avaliação deve ser eficaz, eficiente, viável, útil, realista e precisa, realiza desde 1996 a avaliação dos docentes pelos discentes, como parte de uma avaliação institucional mais ampla.

Segundo Miller (1998, *apud* ALMEIDA, 2002), o corpo docente de uma universidade é a principal fonte do conhecimento acadêmico. A qualidade do ensino é basicamente determinada pelas ações acadêmicas que norteiam o processo de ensino e de aprendizagem. Essas duas afirmações evidenciam que o bom professor é aquele que é conhecedor da sua disciplina e possuidor de habilidades técnicas para tornar o seu ensino mais efetivo. Assim, a universidade precisa de um corpo docente plenamente atualizado nas diversas áreas de conhecimento e possuidores de habilidades técnicas suficientes para intermediar o processo de ensino e de aprendizagem.

Ciente de que a qualidade do ensino é basicamente determinada pela capacidade de seus professores, a avaliação sistemática do seu desempenho é de importância primordial para o melhoramento desses e, como consequência, o dos cursos de graduação. É preciso salientar que, embora os alunos sejam a principal fonte de informação sobre a efetividade do professor, é preciso cruzar os níveis de satisfação desses com a avaliação de colegas professores, bem como a auto-avaliação do professor, de ex-alunos, de coordenadores de curso e de outros, para a obtenção de informações convergentes que possibilitem caminhos para o aprimoramento contínuo do professor.

A avaliação dos docentes pelos discentes, como parte de uma avaliação institucional mais ampla, é um procedimento válido quando a comunidade universitária conscientizar-se que o fim legítimo da avaliação é a busca constante de informações de como se processam o ensino e a aprendizagem no âmbito da sala de aula. Os resultados decorrentes do processo avaliativo são utilizados nas redefinições das metas institucionais, com vistas ao melhoramento contínuo da qualidade do corpo docente e dos cursos de graduação.

A avaliação dos docentes pelos discentes, como parte do processo de Avaliação Institucional defendida entre outros por Carvalho *et al.* (1998), Nunes (2003), Naujorks *et al.* (2003), Silva (2003), Siluk *et al.* (2003), Prestes *et al.* (2003), Duarte *et al.* (2003), Rösler *et al.* (2003), Ramos (2003), Lehmkuhl (2002), Universidade do Oeste de Santa Catarina (2003), parte da hipótese de que existem conjuntos de atributos definidores da boa docência que são distribuídos tecnicamente em um questionário que será valorado pelos alunos. Assim, obtêm-se respostas rápidas, econômicas, confiáveis e válidas sobre como os diversos indicadores da boa docência estão sendo percebidos pelos discentes.

Outro aspecto fundamental é quanto à qualidade do questionário utilizado na avaliação do desempenho do aluno na disciplina. A qualidade das conclusões é decorrente da existência de um conjunto de dimensões confiáveis e válidas da qualidade do ensino ou dos diversos critérios considerados relevantes para o processo de ensino e de aprendizagem. Um instrumento bem elaborado é condição necessária, embora não seja suficiente, para o sucesso de uma pesquisa de avaliação (MOREIRA, 1986). A utilização de instrumento com capacidade técnica bem estabelecida favorece a análise crítica e positiva dos resultados decorrentes do processo avaliativo por parte dos professores, coordenadores de cursos e outros.

Malhotra (2001) define a confiabilidade como o grau em que uma escala produz resultados consistentes, quando se realizam medições repetidas das características em análise. No caso da confiabilidade das avaliações dos alunos, esta é comumente determinada pelo grau de consistência das respostas dos diferentes alunos, numa mesma classe, ao atribuírem o seu nível de satisfação para com os distintos indicadores, destinados a medir determinado conceito do instrumento de avaliação. A validade de um instrumento de medida é o grau com que os indicadores unidimensionais da escala medem com consistência interna as distintas dimensões do instrumento que foram designadas para medir (RAMOS, 1986).

Portanto, para conduzir uma gestão eficaz da docência universitária, é essencial que a tomada de decisões seja sustentada adequadamente por dados oriundos de instrumentos de avaliação dotados de alta qualidade técnica. Sem dados confiáveis, percepções e intuição tornam-se, geralmente, forças dominantes por trás do processo de tomada de decisões em todos os níveis (BURKHALTER, 1993).

Newall (1991, *apud* OWLIA; ASPINWALL, 1996) revisou várias publicações e concluiu que, somente com medidas eficazes é possível avaliar as conquistas provenientes de ações voltadas para a melhoria da qualidade. Isso significa que, sem informações confiáveis e válidas sobre como o sistema está funcionando, não há evidências ou provas de melhorias.

Espera-se, com o apoio das técnicas estatísticas multivariadas, tais como as de análise de componentes principais, análise fatorial exploratória e análise fatorial confirmatória, propor uma metodologia que possibilite a sua utilização na extração de informações relevantes e necessárias para o estabelecimento das propriedades de confiabilidade e de validade do instrumento de avaliação da docência pelos discentes, utilizadas pelas instituições de ensino. Pretende-se submeter os dados provenientes dos processos avaliativos a esta metodologia para diagnosticar com menor erro possível os indicativos de confiabilidade e de validade dos instrumentos de avaliação. O cuidado para com o uso de instrumentos de coleta de dados pertinentes, válidos e confiáveis, a fim de permitir aos professores, coordenadores de curso e outros, a tomada de decisões necessárias para melhorar as ações e resultados refletem o avanço positivo quanto à cultura da avaliação no âmbito da universidade.

1.1 Justificativa e importância do trabalho

Os relatórios decorrentes da avaliação dos docentes pelos discentes das universidades brasileiras ainda são muito questionados pelos professores na atualidade. Alguns argumentam que os resultados informados são inconsistentes devido à falta de maturidade dos discentes para perceber que determinadas ações implementadas pelos docentes fazem-se necessárias para o próprio aprendizado. Acreditam, ainda que os alunos somente terão condições de julgá-los após a inclusão desses no mercado de trabalho e, também, atuarem em áreas correlacionadas com a carreira escolhida. Outros argumentos são que os resultados das avaliações são influenciados por variáveis, tais como o tamanho da turma, dificuldade intrínseca da disciplina, carga elevada de trabalho, falta de motivação dos alunos e outras, e não são controladas no processo de avaliação e crítica dos dados, necessitando, portanto, de controle por parte da equipe de avaliação das instituições de ensino, com vistas a uma pontuação isenta de erros.

Todos esses argumentos são justos e legítimos na ausência de estudos que evidenciem as propriedades fundamentais do instrumento de medida, quais sejam: a sua confiabilidade e a sua validade. O suporte empírico para a validade do instrumento é obtido quando, no processo de medição, as fontes sistemáticas e aleatórias de erro agirem de forma minimizada, ou seja, a medida não é influenciada por variáveis que influenciam nos níveis de satisfação dos alunos, tal como a presença do professor na sala de aula, e nem por variáveis, tais como o estado de humor do aluno ou temperatura ambiente da sala no dia da valoração pelos alunos do instrumento de medida.

Se o instrumento de avaliação da docência pelos discentes providenciar medidas que contemplem médias altas aos professores que alcançarem seus objetivos, isto é, que fizeram com que realmente houvesse aprendizagem e médias baixas para os docentes que ensinaram e/ou exigiram pouco, sendo essa diferença justificada pelos atributos pedagógicos

do bom professor, é possível concluir que o instrumento é válido para medir a qualidade do ensino no âmbito da sala de aula.

Tendo assegurado que existem 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 estabelecidas as propriedades fundamentais do instrumento de medida, quais sejam, a sua confiabilidade e validade.

Esse indicativo é importante para que o professor, ao analisar seu relatório de desempenho na disciplina, esteja confiante de que as informações ali contidas são fidedignas e válidas para indicar as dimensões da qualidade do ensino que devem merecer maior atenção por parte do professor. São igualmente úteis para os gestores da universidade desenvolverem projetos educacionais, visando ao aperfeiçoamento do professor em dimensões nas quais os alunos manifestaram baixo nível de satisfação.

Nesse ponto, existem argumentos suficientes para justificar esta tese que se relaciona diretamente com a qualidade do instrumento de medida usada pelas equipes de avaliação institucional das instituições de ensino superior. Da comprovação das qualidades técnicas do instrumento, decorre maior confiabilidade e validade para os resultados das avaliações, tão necessárias para a manutenção do sucesso de qualquer instituição de ensino.

1.2 Problema e hipóteses básicas da pesquisa

1.2.1 Problema

Diversas são as definições de avaliação que mencionam que as informações decorrentes do processo avaliativo devam apresentar as características de confiabilidade e de validade. Como exemplo, pode-se citar da Conferência Mundial sobre o Ensino Superior (1998, p. 148) que define a avaliação como um “processo de coleta e tratamento de informações pertinentes, válidas e confiáveis, permitindo aos agentes responsáveis tomar as decisões que se impõe, para melhorar as ações e os resultados”.

Embora essas duas propriedades fundamentais das medidas sejam focalizadas nas diversas definições de avaliação, observa-se, nos encontros nacionais ou estaduais que tratam do tema “avaliação”, que as universidades iniciam o processo de avaliação institucional pela avaliação dos docentes pelos discentes, porém não fazem referência quanto à confiabilidade e validade dessas medidas. A questão de validade de conteúdo dos instrumentos é plausível, uma vez que, nas instituições de ensino, os instrumentos são construídos com a colaboração de toda a comunidade acadêmica. Mas, mesmo estando ciente da validade de conteúdo desses instrumentos, não se tem a certeza de que os alunos, ao responderem as questões ali contidas, o fazem consistentemente ou mesmo que esses indicadores considerados relevantes estão associados com o conceito medido. É preciso que os responsáveis pelo setor de avaliação das instituições de ensino investiguem a qualidade de seus dados, utilizando-se de procedimentos empíricos de análise de dados e, desse modo, oferecer maior segurança para os que fazem uso dos resultados decorrentes do processo avaliativo em suas tomadas de decisões. Assim, o problema de pesquisa está relacionado com os indicadores de confiabilidade e de validade do instrumento de avaliação da docência pelos discentes das instituições de ensino.

A falta de uma metodologia específica para a comprovação dessas duas propriedades fundamentais do instrumento avaliativo, torna as questões de confiabilidade e de validade das medidas, pouco discutidas em nível de Brasil. Esse fato é facilmente comprovado pelo número insignificante de publicações que tratam sobre o tema na literatura

brasileira. Devido a essa dificuldade, as instituições de ensino aceitam a hipótese de que os seus instrumentos avaliativos são adequados, e passam a valorá-los pelos alunos.

Portanto, este trabalho pretende dar resposta a seguinte questão: É possível propor uma metodologia que teste as propriedades de confiabilidade e de validade das dimensões subjacentes nos instrumentos de avaliação dos docentes pelos discentes das instituições de ensino?

1.2.2 As hipóteses básicas

Aubrecht (1981) deixa claro que os assuntos básicos em torno do uso do nível de satisfação dos alunos em relação à instrução recebida, estão relacionados com a confiabilidade e a validade do instrumento de avaliação.

Assim, para que as avaliações dos alunos sobre a qualidade do ensino de determinado professor sejam válidas, elas devem medir com consistência interna as diversas dimensões do instrumento, destinadas a medir a qualidade do ensino no âmbito da sala de aula.

Dentro dessas perspectivas, as hipóteses abaixo serão testadas como subsídio à maior confiança por parte daqueles que farão o uso das informações decorrentes do processo de avaliação pelos alunos do desempenho da disciplina.

1) Para os propósitos da avaliação formativa ou somativa:

As avaliações dos docentes pelos discentes sobre as características do ensino daquele professor em particular são confiáveis.

2) Para propósitos da avaliação formativa:

As avaliações dos discentes sobre as características do ensino daquele professor em particular providenciam informações válidas que o auxiliem na melhora contínua da qualidade do seu ensino.

As confirmações dessas análises, além de providenciar seguridade para o uso das diversas informações decorrentes do processo avaliativo, dotarão o instrumento de alta qualidade técnica, uma vez que confiabilidade e validade proporcionam os índices de qualidade desejáveis dos instrumentos de avaliação.

É preciso explicitar 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 em um conjunto de decisões que se apóiam 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.

1.3 Objetivos

1.3.1 Geral

Este trabalho tem como objetivo propor uma metodologia para testar as propriedades de confiabilidade e de validade de instrumentos de avaliação da docência pelos discentes em instituições de ensino.

A metodologia será testada na análise das medidas decorrentes do processo de avaliação de disciplinas, realizada pela Universidade do Sul de Santa Catarina – UNISUL, no

segundo semestre letivo de 2002. Da comprovação das hipóteses enunciadas para dar respostas aos objetivos específicos citados abaixo, decorem os indicativos de confiabilidade e validade das medidas.

1.3.2 Específicos

- Examinar a estrutura fatorial do instrumento, utilizando-se das técnicas de análise fatorial exploratória e confirmatória (validade de construto).
- Determinar a confiabilidade e validade dos indicadores e das dimensões subjacentes ao instrumento de avaliação, utilizando-se dos dados “geral da UNISUL” e estratificados segundo o campus.
- Estimar os efeitos e a importância relativa das dimensões da escala sobre os indicadores considerados relevantes pela instituição.
- Determinar a capacidade do questionário para o estabelecimento de tipologias ou perfis de qualidade docente.
- Determinar o grau de correlação entre as dimensões da escala e os indicadores considerados relevantes pela instituição.

1.4 Estrutura do trabalho

O trabalho está estruturado em nove capítulos distintos. O segundo capítulo trata da multidimensionalidade dos instrumentos de avaliação do docente através do discente, no qual são abordadas as dimensões de alguns instrumentos de avaliação utilizados em universidades americanas, espanholas e brasileiras.

O terceiro capítulo apresenta os aspectos teóricos da confiabilidade. Nele, são discutidas as fontes aleatórias e a sistemática de erros, bem como a sua influência nas medidas. São abordadas a confiabilidade global e a confiabilidade por dimensões específicas de alguns instrumentos de avaliação da docência amplamente discutidos na literatura. Discute a questão da estabilidade das avaliações dos alunos e relata algumas pesquisas que trazem à tona que as ações do professor é que discrimina um professor do outro.

O capítulo quatro trata dos aspectos teóricos da validade, com ênfase na validação de construto, uma vez que esse tipo de validade engloba as validades de conteúdo e de critério. Discute alguns métodos para acessar a validade de construto e relata que existem dimensões de instrumentos de avaliação da docência que refletem a efetividade do ensino. Traz, também, o relato de algumas pesquisas que tratam da influência de variáveis indesejáveis nas avaliações dos discentes.

O capítulo cinco trata dos procedimentos metodológicos do processo de análise fatorial exploratório. É apresentado como importante técnica para evidenciar as dimensões subjacentes aos dados empíricos e de grande importância para ajudar o pesquisador com bons conhecimentos teóricos, na elaboração ou reformulação dos conceitos.

O capítulo seis destina-se ao método contemporâneo de análise fatorial confirmatória, para acessar a confiabilidade e a validade de instrumentos de avaliação. Apresenta todo o enfoque metodológico para testar as hipóteses decorrentes dos estudos exploratórios.

O capítulo sete descreve como a metodologia que será aplicada para validar as dimensões subjacentes ao instrumento “Avaliação pelos alunos do Desempenho na Disciplina” utilizados pela equipe de avaliação e clima organizacional da UNISUL. Informa que o método da pesquisa é o hipotético-dedutivo e descreve as hipóteses e sub-hipóteses

necessárias, para evidenciar as propriedades de confiabilidade e validade das dimensões destinadas a refletirem a qualidade do ensino no âmbito da sala de aula.

O capítulo oito está reservado à apresentação dos resultados dos testes.

Na seqüência destacam-se as considerações finais da pesquisa realizada. Por fim, as referências bibliográficas consultadas e citadas e os anexos contendo o instrumento de avaliação da docência da UNISUL e das matrizes utilizadas para a comprovação das propriedades estatísticas de confiabilidade e validade do instrumento de avaliação em análise.

2 MULTIDIMENSIONALIDADE DOS INSTRUMENTOS DE AVALIAÇÃO DA DOCÊNCIA

2.1 Introdução

A avaliação dos professores pelos alunos, como parte de uma avaliação institucional mais ampla, tem como finalidade alimentar os professores e gestores da instituição de ensino com informações relevantes, confiáveis e válidas de como está ocorrendo o processo de ensino/aprendizagem, no âmbito da sala de aula. O objetivo fundamental é fornecer informações que possibilitem a reflexão por parte dos professores de sua atuação, bem como o estabelecimento de programas que possibilitem a formação continuada do docente.

O instrumento mais utilizado para recolher a opinião dos discentes é o questionário. A maioria deles solicita que os alunos manifestem o seu nível de satisfação, geralmente numa escala tipo *likert* de quatro a sete categorias de respostas, em relação aos aspectos considerados relevantes para uma boa docência.

Tejedor (1996) reconhece que o critério de referência dos alunos seja o mais utilizado, mas cita que nenhum critério seja totalmente eficaz. É preciso combinar as informações dos alunos com as de outras fontes, tais como auto-avaliação do professor, de colegas professores e outros, de modo a acumular informações convergentes que apontem caminhos para a melhoria da qualidade do ensino.

Quanto ao uso dos resultados das avaliações pelos docentes, muitos discutem a falta de validade dessas avaliações argumentando que os alunos são imaturos e que só depois de formados e trabalhando na sua área de formação, poderiam avaliar com segurança o desempenho do professor na sua disciplina. Mas, inúmeras pesquisas e comentários de

especialista na área, atestam que os alunos podem fazer julgamentos consistentes sobre a qualidade do corpo docente e sobre o ensino recebido. Nesse sentido, para Castro (1992), os discentes reconhecem a dedicação dos docentes, a sua motivação para o ensino da disciplina, a excelência da didática utilizada, e outras. Scriven¹ (1995) coloca que os alunos estão numa posição única para avaliar o seu conhecimento crescente nas disciplinas matriculadas no semestre, por isso podem observar e avaliar aspectos importantes, tais como a pontualidade do professor, a legibilidade da escrita no quadro, identificar e avaliar o grau de entusiasmo do professor durante as aulas e, também, a forma como os professores se relacionam com a turma. Todas estas características são relevantes para alcançar um ensino competente. Portanto, parece ser os alunos os indivíduos com mais condições de efetuarem a avaliação do professor.

Evidências empíricas, sobre a maturidade das opiniões dos alunos, foram constatadas por Overall e Marsh (1980). Eles estudaram os níveis de satisfação dos alunos quanto ao desempenho dos professores, no âmbito da sala de aula, no final de cem cursos. Uma vez mais os mesmos alunos foram convidados a expor o nível de satisfação com os mesmos professores, um ano após a sua formatura. Da análise, resultou um coeficiente de estabilidade médio de 0,83. Em face à alta correlação reportada, os autores concluíram que os alunos conservam as mesmas opiniões sobre a qualidade do ensino dos professores ao longo do tempo.

Um outro estudo experimental foi realizado por Howard *et al.* (1985), com o objetivo de discutir qual é o melhor estimador da efetividade do professor no âmbito da sala de aula, quais sejam: avaliações dos professores, avaliações dos alunos, as avaliações de ex-alunos. O estudo revelou que todos esses avaliadores da eficácia do professor são úteis, mas que, em particular, as avaliações feitas pelos alunos e ex-alunos apresentam maiores coeficientes de validade.

¹ Todas as citações em língua não vernácula foram traduzidas pelo autor.

2.2 Utilidade das avaliações dos alunos

Existem evidências certificadas substanciais como também evidências experimentais, apoiando a utilidade das avaliações dos docentes por parte dos discentes (CENTRA, 1993; MARSH e DUNKIN, 1992; MARSH e ROCHE, 1994; COHEN, 1980).

Costin *et al.* (1971) e Marsh (1984) apresentam que as avaliações dos alunos podem ser úteis para:

- 1) providenciar *feedback* para o docente melhorar o seu ensino, uma vez que nem todos os alunos são capazes de efetuar uma avaliação do docente na base face a face;
- 2) providenciar uma maneira através da qual um membro docente poderia, se desejasse, demonstrar sua eficácia no ensino de graduação para aqueles que têm expressado um interesse em avaliar esse parâmetro para uma promoção;
- 3) informação para os alunos usarem na seleção de disciplinas e professores;
- 4) providenciar informações para a coordenação do curso sobre dimensões de relativa força ou fraqueza no ensino de graduação; e
- 5) providenciar uma base de dados para a pesquisa sobre o ensino.

Quanto ao uso das avaliações dos alunos pelos docentes para o melhoramento da qualidade do ensino, a meta-análise realizada por Cohen (1980), relata que os professores melhoram o desempenho da docência em função do *feedback* realizado pelos discentes. Porém, observou-se que os professores que receberam a sua avaliação e o acompanhamento da assistência de um consultor educacional, obtiveram, na avaliação do final do semestre, 16 pontos percentuais, em média, superior aos professores que receberam a avaliação no final do semestre sem ter tido o acompanhamento de um consultor educacional.

Segundo Aleamoni (1978), Cohen (1980), Marsh e Overal (1979) e Mckeachie *et al.* (1980) (*apud* MCKEACHIE, 1997), os professores que recebem orientações de um

consultor educacional, designado para orientar os professores quanto ao melhor uso dos resultados das avaliações feitas pelos alunos, conseguem melhorar com índices razoáveis o nível de qualidade do seu ensino. Para Lawall (1998), o papel do consultor é o de direcionar a escolha de estratégias de melhorias, assegurando a implementação de estratégias selecionadas e direcionando o processo de reunir retorno adicional em relação àquelas estratégias implementadas pelo professor.

Para Marsh e Roche (1997), o retorno das avaliações dos alunos por dimensões específicas da qualidade do ensino é o meio mais útil de aumentar a sua utilidade, dado que ajuda o professor a diagnosticar pontos fracos evidenciados pelo processo avaliativo.

A validade do uso das avaliações dos alunos, para providenciar algum mérito ao docente, são fontes de dados válidos para a eficácia do ensino do professor, mas o seu uso deveria ser melhorado, segundo Mckeachie (1997). De acordo com esse autor, comparações de avaliações de professores de classes diferentes são duvidosas, não apenas por causa da diferença entre classes de alunos, como também diferenças entre objetivos, métodos de ensino, conteúdo e outras. Para D'Apollonia e Abrami (1997), as avaliações dos alunos a respeito da docência do professor providenciam informações válidas para atribuição de algum mérito ao docente, porém elas não deveriam ser a única fonte de informação, e nem deveriam ser interpretadas demasiadamente.

Quando as avaliações dos alunos forem usadas para a promoção do corpo docente, d'Apollonia e Abrami (1997) argumentam que as dimensões da qualidade do ensino são correlacionadas e que elas podem ser representadas através de um único fator hierárquico-Habilidade Instrucional Geral. Assim, um escore único, representando as habilidades instrucionais gerais do professor, deveria ser usadas para avaliação somativa.

Segundo Cashin (1995), os assuntos básicos em torno do uso do nível de satisfação dos alunos em relação à qualidade do ensino recebido estão relacionados com a

confiabilidade e a validade das avaliações estudantis. Portanto, todos os benefícios citados por Costin *et al.* (1971) e Marsh (1984) podem existir apenas na extensão em que as opiniões do aluno representam avaliações confiáveis e válidas sobre a instrução recebida em sala de aula.

Como pode ser visto, a utilidade das avaliações estudantis aumentam desde que, na universidade, exista um programa institucional alimentado pelos resultados das avaliações dos professores pelos discentes, cuidadosamente planejado para elevar o nível da qualidade do ensino dos professores e, conseqüentemente, dos cursos de graduação. Somente poderão avaliar os méritos do programa os professores que se dedicam na busca constante de estratégias para melhorar o nível de aprendizado dos alunos.

Também é essencial o apoio da administração universitária. Uma administração universitária preocupada com o aprendizado dos alunos será percebida pelos docentes, que se esforçarão para melhorar o nível da qualidade de seus ensinamentos e pelos discentes, reconhecendo que estudam em uma universidade que se preocupa em formar profissionais qualificados para o mercado de trabalho.

Os resultados das avaliações dos alunos, como um indicador do ensino eficaz, é uma base de dados útil, tanto a avaliação formativa como somativa do professor universitário. Entende-se por avaliação formativa aquela planejada para fornecer informações úteis à auto-análise dos professores e gestores com vistas ao melhoramento da qualidade do ensino em todos os níveis. A avaliação somativa é realizada com o objetivo de fornecer informações fidedignas que possibilitem decisões sobre o futuro do professor, tais como efetivação, promoção e outra.

2.3 Dimensionalidade, confiabilidade e validade

Os assuntos básicos em torno do instrumento de avaliação da docência pelos alunos estão relacionados com a sua dimensionalidade, confiabilidade e validade (CASHIN, 1995).

Para Hair *et al.* (1999), a dimensionalidade está associada ao número de características de um objeto. No caso de instrumento de avaliação dos professores, a dimensionalidade refere-se ao número de dimensões do instrumento de avaliação, considerado relevante para a efetividade do ensino. O instrumento de avaliação de disciplinas da UNISUL, *a priori*, foi concebido com oito dimensões, consideradas pela comunidade acadêmica – administrativo, professores, alunos, especialista em medição educacional como atributos necessários para a garantia da qualidade do ensino. As dimensões do instrumento são: exigência, avaliação, conteúdo, método, relacionamento professor aluno, produtividade, ética e relação professor curso.

De acordo com Malhotra (2001), a confiabilidade é o grau em que uma escala produz resultados consistentes, quando se realizam medições repetidas das características em análise. No caso da confiabilidade das avaliações dos alunos, esta é comumente determinada pelo grau de consistência das respostas dos diferentes alunos, numa mesma classe, ao atribuírem o seu nível de satisfação para com os distintos indicadores, destinados a medir determinado conceito do instrumento de avaliação. O termo indicadores é utilizado para designar o conjunto de variáveis que representam o conceito medido (HAIR *et al.* 1999).

A validade de um instrumento de medida é o grau com que os indicadores unidimensionais da escala medem com consistência interna as distintas dimensões do instrumento que foram designadas para medir (RAMOS, 1986). Quanto ao instrumento de avaliação utilizado pela UNISUL, esse será válido se as nove dimensões já citadas forem

extraídas dos dados empíricos por procedimentos de análises fatorial exploratória e confirmatória. Marsh (1984) diz que as avaliações dos alunos constituem uma medida da qualidade do ensino recebido, sendo difíceis de validar, devido à existência de múltiplos critérios que definem o ensino eficaz.

Estudos exploratórios são aqueles que definem as possíveis relações de uma forma mais geral, enquanto os estudos confirmatórios testam a veracidade dos relacionamentos especificados (HAIR *et al.* 1999).

Churchill (1979), coloca que uma medida tem validade, quando as diferenças entre os valores observados refletirem diferenças verdadeiras sobre as características que se pretende medir e não sobre outros fatores. Assim, a validade de uma escala pode ser definida como o âmbito em que as diferenças entre pontuações observadas da escala, com relação ao desempenho dos professores, sejam explicadas pelas verdadeiras características que distinguem um bom professor de outro e não pela influência de outras variáveis.

2.4 Paradigmas norteadores da avaliação da docência

Ao se considerar que o professor universitário é a principal fonte de conhecimento acadêmico, que desempenha uma função prioritariamente técnica, uma vez que o seu desempenho no ensino depende de sua habilidade para relacionar-se com a classe de alunos, e que, também, um adequado nível de relacionamento é alcançado quando o docente consegue integrar de forma inteligente os conteúdos a serem trabalhados na disciplina e a técnica, sabe-se que não é qualquer profissional que se designe a ministrar determinada disciplina e que o faz com competência. É preciso que o professor seja preparado continuamente para o exercício da docência, para atender, adequadamente, as exigências dos novos alunos que semestralmente chegam às universidades.

Dentro dessa perspectiva, segundo Doyle (1977, *apud* TEJEDOR *et al.* 1990) derivam-se três paradigmas que inspiram as diferentes estratégias de avaliação do ensino, quais sejam: o paradigma processo - produto, o paradigma mediador e o paradigma interativo. Borich (1986 *apud* SUCH, 2003) acrescenta aos paradigmas anteriores, o paradigma prognóstico - produto. Em síntese, esses paradigmas caracterizam-se pelos seguintes aspectos:

- **paradigma prognóstico-produto** - Segundo esse paradigma, a competência do professor está baseada nas características físicas e psicológicas, definidoras de sua personalidade. São indicadores desse paradigma: o entusiasmo, o compromisso, o domínio da matéria, a habilidade de comunicação, a clareza expositiva, o pensamento lógico, a capacidade pedagógica. Esses são os principais atributos do professor necessários para uma boa docência, ou seja, essas variáveis causam o aprendizado dos alunos.
- **Paradigma processo-produto** - Trata de estabelecer as relações entre o processo “conduta do professor” e os resultados alcançados pelos alunos. Esse paradigma entende que o professor é a causa imediata do aprendizado do aluno. São variáveis desse paradigma: ordem expositiva, seqüência lógica, relacionamentos entre os temas, elaboração de sínteses, avaliação do aprendizado, uso dos recursos de apoio pedagógico, relacionamentos com os alunos. Esse paradigma engloba o anterior.
- **paradigmas dos processos mediadores** - Esse paradigma inclui todos os aspectos do paradigma anterior. Reconhece e utiliza variável adicional relativa ao contexto escolar e social do aluno, bem como a experiência do discente e do docente. São variáveis desse paradigma: a cultura geral do professor, capacidade técnica para o uso dos recursos instrucionais, conhecimento técnicos do professor sobre aprendizagem, nível de conhecimentos dos discentes, consciência metacognitiva e outras. Esse paradigma sustenta a hipótese de que o professor não influi de modo direto no aprendizado do aluno, e sim em

alguns processos que constituem a atividade cognitiva desse, ocorrendo, assim, o seu aprendizado.

- **paradigma interativo** - Segundo esse paradigma, a eficácia do professor para com o aprendizado dos discentes ocorre em contexto específico. Visualiza e reconhece a complexidade das relações e trocas que ocorrem durante o processo de ensino e de aprendizagem. São variáveis desse paradigma: aspectos de desenvolvimento físico, psicológico e social dos alunos, habilidades de comunicação do docente e dos discentes; uso de estratégias diversas para a solução de problemas, avaliação centrada na compreensão e do domínio do tema.

Segundo Tejedor *et al.* (1988), não existe acordo entre os pesquisadores sobre quais variáveis do paradigma processo-produto relacionam-se melhor com o aprendizado do aluno, e que ainda, esse paradigma é ou não suficiente para consolidar um instrumento de avaliação da docência pelos alunos, uma vez que o mesmo responsabiliza totalmente o professor pelo seu aprendizado. É necessário obter medidas indiretas do desempenho dos docentes, gerando um modelo multidimensional de avaliação, integrando, quando possível dimensões específicas a todos esses paradigmas.

2.5 Teorias implícitas dos alunos

Durante as etapas de identificação das dimensões dos comportamentos dos professores que mais se identificam com o ensino de qualidade, as técnicas exploratórias e confirmatórias de análises de dados ajudam o pesquisador a estabelecer a ponte entre as variáveis da qualidade do ensino e o suporte empírico da dimensionalidade dos comportamentos dos professores.

Withely e Doyle (1978), chamam a atenção quanto a dois aspectos relacionados com a estrutura fatorial resultante da análise empírica: a estabilidade da estrutura fatorial e até que ponto as estruturas fatoriais representam as teorias implícitas dos alunos.

A estabilidade da estrutura fatorial é alcançada, quando o mesmo instrumento for submetido à valoração estudantil em contextos distintos e por pesquisadores diversos e a análise das valorações dos estudantes, mediante o uso de análise fatorial exploratória e confirmatória, convergirem quanto ao número e significado das dimensões.

Teorias implícitas dos alunos diz respeito ao conjunto de idéias ou crenças que os alunos internalizaram na sua infância e adolescência sobre o conceito do bom professor, já que, até então, cada um deles teve vários tipos de professores e de ensino e trazem para a universidade. Essas teorias implícitas sobre o conceito de bom professor que os alunos trazem para a universidade facultam-lhes transformar a avaliação numa simples comparação entre o professor presente e o professor ideal.

A pesquisa de Doyle e Whitely (1978) analisou quatro amostras de estudantes que avaliaram diferentes tipos de professores. Os autores concluíram, através de estruturas fatoriais separadas, que os estudantes mantêm teorias implícitas comuns sobre as características do ensino e que os dados sustentam a hipótese de que os fatores do ensino efetivo são altamente estáveis em relação às distintas disciplinas e tipos de professores.

A similaridade das duas estruturas fatoriais dos instrumentos de avaliação dos professores e alunos, comprovados por técnicas analíticas de análise de dados, fornece o suporte empírico para a validade convergente das dimensões da avaliação.

Whitley e Doyle (1978) sugerem que as teorias implícitas dos alunos sobre as habilidades e comportamentos dos docentes não invalidam as avaliações dos alunos, pois elas revelam experiências anteriores vivenciadas pelos alunos e trazem-nas para a universidade, sendo, assim, perfeitamente válidas.

Para Marsh (1984), embora alunos e professores possam ter teorias implícitas similares, os professores são os únicos capazes de observar os seus próprios comportamentos e têm pouca necessidade para confiarem em teorias implícitas, na formação de suas auto-avaliação.

2.6 Dimensionalidade dos instrumentos de avaliação da docência

De acordo com Marsh (1982), o instrumento de avaliação da qualidade do ensino a ser valorado pelos alunos deve ser multidimensional, ou seja, deve contemplar todas as dimensões do processo de ensino e de aprendizagem que conduzem para a eficácia do ensino.

As dimensões da qualidade do ensino, no âmbito da sala de aula, foram obtidas pela maioria dos pesquisadores através do recurso analítico de análise fatorial exploratória, aplicado aos dados empíricos. Essa prática foi a mais empregada até o final dos anos 70. Com o avanço das novas tecnologias de análise, a partir dos anos 80, a técnica de análise fatorial confirmatória, passou a ser utilizada, no âmbito educacional, com vistas a confirmar as dimensões decorrentes do processo exploratório. Os estudos confirmatórios são utilizados, visando o estabelecimento pleno da validade de construto, uma vez que, na maioria das vezes, os instrumentos são designados para medirem vários conceitos simultaneamente.

Para Ramos (1986), validade de construto é o tipo de validade que avalia o grau em que uma escala mede o conceito para o qual foi designada a medir. Em se tratando de instrumento multidimensional a validade de construto indica o grau em que os distintos indicadores unidimensionais do instrumento medem consistentemente os diferentes conceitos para os quais foram designados a medir.

Marsh (1984), cita que muitos dos instrumentos de avaliação são construídos sem reconhecer essa multidimensionalidade e, conseqüentemente, não há base para saber o que se

está medindo. Se o instrumento contém grupos separados de indicadores relacionados com a qualidade do ensino, derivados de uma análise lógica do conteúdo, e se esses relacionamentos demonstrados por procedimentos empíricos, evidenciarem que indicadores do mesmo grupo de fato medem em separado traços distinguíveis da qualidade do ensino do professor, então é possível interpretar o que está sendo medido.

Segundo Marsh (1984), existem pesquisadores que não conseguem evidenciar, por meio de procedimentos empíricos, a multidimensionalidade da qualidade do ensino, através de seus instrumentos e pregam que uma avaliação globalizada da qualidade do ensino seja uma medida mais válida. Segundo esse autor, a falta de evidências da multidimensionalidade do instrumento de avaliação enfraquece a sua utilidade como *feedback* para a avaliação formativa do docente.

A elaboração de instrumentos unidimensional para a medição da qualidade do ensino, não é adequada pelas seguintes razões:

- 1- existem muitos indicadores possíveis do ensino eficaz e muito usos possíveis para as avaliações dos alunos. Para Marsh e Overall (1980, *apud* MARSH, 1984), o componente que é mais válido dependerá dos critérios a serem considerados pelas instituições de ensino;
- 2- pesquisas relatam que as dimensões do instrumento de avaliação dos alunos estão mais altamente correlacionadas com outros critérios da qualidade do ensino. Por exemplo, Cohen (1981) reportou uma correlação de 0,47 entre as notas dos alunos na prova externa (prova elaborada por professores da área, mas que não ministram aulas em nenhuma seção), como medida do aprendizado do aluno, com os níveis de satisfação destes com a dimensão “estrutura do professor”. Capelleras e Veciana (2000) encontraram correlação de 0,58 entre a variável critério “qualidade global da universidade” e a dimensão “atitudes e comportamento dos professores”;

- 3- a influência de uma variedade de variáveis supostas a enviesarem as avaliações dos alunos é mais fácil de interpretar quando as dimensões dos instrumentos forem correlacionadas com essas variáveis enviesantes. Por exemplo, Marsh (1984) reporta que a correlação entre a variável “tamanho da classe” está negativamente correlacionada com duas dimensões do instrumento SEEQ: “interação do professor com o grupo” e “relacionamento individual” e não correlacionada com as demais dimensões;
- 4- quando as avaliações dos alunos são utilizadas como diagnóstico para os docentes, a sua utilidade aumenta quando os retornos das avaliações são disponibilizados segundo as dimensões. Por exemplo, mais rapidamente o professor pode constatar se os seus alunos estão mais satisfeitos com a dimensão “relacionamento” ou com a dimensão “avaliação”; e
- 5- nenhum docente satisfaz plenamente a todos os discentes em todas as dimensões da qualidade do ensino. Por exemplo, um professor pode apresentar uma aula muito bem estruturada, mas pode não conseguir motivar suficientemente os alunos para o aprendizado do tema.

As investigações de Marsh (1982), Tejedor *et al.* (1988), Frey *et al.* (1975), Ramos (1997), Hildebrand *et al.* (1971), Warrington (1973) mostram um amplo nível de acordo sobre a multidimensionalidade, associado às diversas características de um ensino adequado e que os alunos através dos seus níveis de satisfação para com os distintos indicadores do instrumento, reconhecem a multidimensionalidade do ensino, ou seja, os alunos podem identificar as dimensões mais importantes de um ensino efetivo.

Embora diversas pesquisas demonstrem que um ensino de qualidade é um construto multidimensional, D’Apollonia e Abrami (1997) e McKeachie (1997) reconhecem a multidimensionalidade do ensino, porém relatam que a avaliação contém um amplo fator global da qualidade do ensino, que determina as várias dimensões inferiores altamente correlacionadas.

A primeira escala publicada, destina-se à avaliação do ensino e dos professores pelos alunos, foi a escala “Purdue Rating Scale”, desenvolvida por Smalzried e Remmers (1943), com dez itens que indicam duas dimensões que diferenciam os professores, quais sejam: relacionamento dos professores com os alunos e a maturidade profissional dos docentes. Essa escala, desde a sua publicação, tem sido motivo de preocupação e de atenção por parte de muitos pesquisadores. Os trabalhos de Herbert W. Marsh, Philip C. Abrami, Sylvia d’Apollonia, William E. Cashin, Anthony G. Greenwal, Gerald M. Gillmore e outros confirmam uma importante linha de pesquisa relacionada com a utilização dos níveis de satisfação dos alunos com respeito a seus professores.

O instrumento de avaliação de Marsh (1982), denominado “Students Evaluations of Educational Quality” (SEEQ), foi testado em vários países por diversos colaboradores, chegando à conclusão de que se trata de um instrumento multidimensional válido para a avaliação do desempenho docente.

Afirma-se, dessa forma, que é o instrumento de avaliação da docência que, até a presente qualificação, apresentou mais provas de que se trata de um instrumento confiável e válido.

São apresentadas, a seguir, as dimensões obtidas em alguns instrumentos de avaliação da docência, e que demonstram evidências empíricas quanto à confiabilidade e à validade em universidades brasileiras, americanas e espanholas.

No Brasil, Moreira (1986) submeteu um formulário reduzido de treze itens, para a avaliação dos docentes pelos discentes da Faculdade de Administração de São Paulo. Foram avaliados 72 professores, possibilitando obter 106 combinações disciplina-classe. Os dados foram submetidos à análise fatorial, resultando um modelo de seis dimensões que explicam 76,1% da variância total da escala. Essas dimensões foram rotuladas respectivamente como: 1) natureza cognitiva ou natureza afetiva; 2) condição de classe; 3) conteúdo; 4) necessidades

profissionais; 5) caráter de mais fácil ou de mais difícil da disciplina; e 6) recursos. O mesmo instrumento foi aplicado em 1985 com o objetivo de investigar a estabilidade fatorial. A comparação da estrutura fatorial obtida em 1982 com a estrutura fatorial obtida em 1985 revelou que as dimensões são apenas parcialmente comparáveis.

Pasquali (1984) submeteu um formulário com 92 indicadores a uma amostra de 354 estudantes de cursos superiores do Distrito Federal. O estudo psicométrico do instrumento possibilitou que os 92 indicadores fossem reduzidos para 71 indicadores. Esses indicadores submetidos à análise fatorial, resultaram em 7 dimensões que explicavam aproximadamente 48,8% da variância total da escala. As dimensões foram interpretadas, respectivamente, como: 1) o grande educador; 2) desrespeito; 3) laboratório; 4) mestre; 5) compreensão da disciplina; 6) respeito por conteúdo e aluno; e 7) imparcialidade com conteúdo e aluno.

O questionário de avaliação da docência universitária, proposto por Tejedor *et al.* (1988), estruturado com 49 itens, foi valorado pelos discentes da Universidade de Santiago, com o objetivo de medir a percepção dos discentes sobre a qualidade do ensino, no âmbito da universidade. Foram recolhidos 44323 questionários para toda a universidade. Os dados foram submetidos à análise fatorial ortogonal e oblíqua, resultando uma estrutura fatorial com oito dimensões que explicavam 67,32% da variância total da escala. As dimensões foram descritas como: 1) clareza, organização, preocupação com os trabalhos; 2) interação com os alunos; 3) avaliação adequada; 4) práticas e material de apoio; 5) desenvolvimento do programa; 6) dedicação; 7) a disciplina e o plano de estudo; e 8) atualização informática.

O instrumento de avaliação desenvolvido por Hildebrand *et al.* (1971) é composto por 40, itens considerados essenciais para o aprendizado efetivo. Estudos fatoriais sobre os dados de avaliação de 338 professores resultaram em 5 dimensões da qualidade do ensino,

quais sejam: 1) aspectos analítico-sintético; 2) organização ou clareza; 3) interação com a classe; 4) interação professor/aluno individual; e 5) dinamismo ou entusiasmo do professor.

Trent e Cohen (1973, *apud* TEJEDOR *et al.*, 1988) compararam vários estudos, estabelecendo as dimensões gerais que definem o professor efetivo e concluíram que os cinco fatores principais são: 1) estrutura e clareza da explicação; 2) estímulo da discussão em classe e apresentação dos diversos pontos de vista; 3) estimulação dos interesses e motivação dos estudantes; 4) manifestação de atenção e interesse pelos alunos; e 5) manifestação de entusiasmo.

O questionário de avaliação de Blanco e Paz (1988), composta por 27 itens, foi aplicado aos estudantes da Universidade de Cantabria, objetivando medir a percepção dos estudantes sobre a qualidade do ensino dos professores dessa universidade. Foram recolhidos 13846 questionários para o total da universidade. Para o total da universidade que responderam ao instrumento, três dimensões que explicam 79,9% da variabilidade total da escala foram designadas para medir a qualidade do ensino no âmbito da sala de aula. Quais sejam: 1) assistência à classe/organização; 2) interação; e 3) domínio/entusiasmo do professor. A dimensão dois, denominada “interação”, manteve-se em todas as análises fatoriais, enquanto as demais foram interpretadas com significados distintos.

Com o objetivo de avaliar a qualidade dos docentes da universidade Complutense de Madrid, Sanchez (1988) utilizou um instrumento de avaliação composto por 27 itens. O instrumento foi preenchido por 1156 alunos. A análise da multidimensionalidade resultou em duas dimensões da qualidade do ensino universitário. As dimensões foram denominadas: 1) competência docente e 2) estimulação/motivação do aluno.

Cajide *et al.* (1992, *apud* VAL, 1994) relata que as dimensões da qualidade do docente percebidas pelos alunos da universidade de Santiago são as seguintes: 1) domínio da

disciplina (competência do professor); 2) interação do professor com os alunos; 3) desenvolvimento do programa; 4) práticas e recursos utilizados; e 5) avaliação.

O formulário de avaliação da qualidade educacional pelos alunos proposto por Marsh (1982), apresenta uma estrutura fatorial com nove dimensões bem consistentes da qualidade do ensino. As nove dimensões foram denominados, respectivamente, como: 1) interação em grupo; 2) aprendizado/valor; 3) trabalho/dificuldade; 4) exames/grau; 5) relacionamento individual; 6) organização/clareza; 7) entusiasmo; 8) aprofundamento dos temas; e 9) tarefas/leituras.

O instrumento de avaliação de docência universitária, proposto por Frey, Leonard e Beatty (1975), denominado de *ENDEAVOR*, é composto por 21 itens. Foi testado na universidade de Northwestern e de Purdue com, respectivamente, 421 e 218 discentes das disciplinas de cálculo e psicologia educacional. As análises fatoriais, para as avaliações de ambas universidades, resultaram em 7 dimensões que oportunizaram a leitura do mesmo conceito. Esses fatores ou dimensões foram rotulados como: 1) discussão em classe; 2) conquistas dos alunos; 3) carga de trabalho; 4) forma de avaliar 5) atenção pessoal; 6) clareza da aula; e 7) organização/planejamento.

Com o objetivo de determinar as dimensões básicas do comportamento do ensino, Umble e Whitten (1977) submeteram um instrumento de avaliação dos docentes, contendo 60 pares bipolares das características dos docentes, a 212 estudantes de negócios da Universidade da Geórgia. Os resultados das valorações estudantis foram submetidos à análise fatorial com rotação ortogonal, resultando seis dimensões que explicavam 48% da variância total da escala. Essas dimensões foram, respectivamente, discutidas como: 1) competência profissional do professor; 2) dominador-dificuldade; 3) comportamento profissional; 4) personalidade na classe; 5) receptividade; e, 6) motivação/atividade. Uma Segunda amostra de 205 estudantes de negócio da mesma universidade foi utilizada para investigar a importância

relativa de cada uma das seis dimensões identificadas. Competência profissional e comportamento profissional foram os fatores facilmente identificados como os mais importantes, enquanto a dimensão “dominador – dificuldade” foi descrita como a menos importante.

Para Aparicio, Tejedor e San Martín (1982, *apud* SANCHEZ, 1985), o questionário de avaliação da docência pelos alunos apresenta os seguintes fatores ou dimensões de qualidade do ensino: 1) entusiasmo/domínio da disciplina; 2) importância da disciplina; 3) organização/claridade/preocupação com o ensino; 4) interesse pelo aluno individual; 5) interação com o grupo e preocupação com os resultados; 6) cumprimento do professor; 7) gentileza do professor; 8) exames; e, 9) material/trabalho /prática.

O instrumento de avaliação do professor pelos alunos de Sherman e Blackburn (1975), na sua forma final, é composto por 22 itens de adjetivos bipolares, tipo *Osgood*. A evidência empírica da dimensionalidade resultou em 4 dimensões que explicam 70% da variância total da escala. Essas dimensões foram rotuladas como: 1) dinamismo do professor; 2) pragmatismo; 3) sociabilidade; e, 4) competência.

O instrumento *SIRS* do Estado de Michigan, desenvolvido por Warrington (1973), é composto por 21 itens que descrevem as características do professor em sala de aula, quatro itens caracterizam os alunos e três itens sobre o professor/laboratório. As dimensões do Sistema de Avaliação Institucional dos Estudantes (*SIRS*) são: 1) envolvimento do professor; 2) interesse e *performance* do estudante; 3) interação professor-aluno; 4) demandas de esforços dos alunos; e, 5) organização da disciplina.

Preocupado com a qualidade do ensino em educação superior, Ramsden (1991) apresenta o *CEQ* (Questionário de Experiência do Curso) como um instrumento confiável, válido e adequado para discriminar com inteligibilidade a *performance* dos cursos pelos estudantes universitários. As cinco dimensões do *CEQ* são as seguintes:

- 1- ensino adequado - Correta orientação sobre a dinâmica do aprendizado do aluno;
- 2- clareza dos objetivos do programa - Os discentes possuem clareza do rumo do curso e do que o curso espera do aluno neste curso;
- 3- adequada carga de trabalho;
- 4- adequada avaliação; e
- 5- estímulo e motivação do aluno em sua responsabilidade - Os alunos são motivados a escolherem os trabalhos, numa variedade ampla de possibilidades.

Ramos (1997), utilizou a abordagem de validação de construto para a estimação de um modelo de avaliação formativa do professor universitário em um centro universitário da Universidade Complutense de Madrid. Seis modelos alternativos foram estimados, utilizando o *LISREL* (Linear Structural Relations) como modelo de análise. As medidas de ajuste absoluto e de ajuste incrementais estabeleceram que o modelo com cinco fatores latentes, correlacionados, seja o mais adequado para medir a competência do docente universitário. As dimensões do instrumento são: 1) programação/organização do ensino; 2) domínio do conteúdo/clareza expositivo; 3) motivação do aprendizado/incremento do interesse do aluno; 4) interação com a classe; 5) atenção individual ao aluno; e, 6) avaliação/exames.

O formulário de avaliação dos docentes pelos discentes, proposto por Macedo (2001), é composto por 16 itens representativos da qualidade do ensino. Os itens foram agrupados em três dimensões teóricas. Cada qual representa um objetivo a ser atingido pelo docente, são elas: 1) domínio técnico científico; 2) habilidade de ensino; e, 3) desempenho pessoal e situacional. O instrumento foi validado com 2980 alunos, os quais valoraram 419 disciplinas, totalizando 175 professores. A média para cada item da escala foi processada e categorizada em percentil. Essas categorias foram submetidas à análise fatorial de correspondência múltipla, possibilitando discriminar diferentes níveis de desempenho docente

quanto às seguintes variáveis de pano de fundo: valor médio na dimensão “desempenho”, a carga horária na instituição, tempo de admissão e a titulação do docente.

Barton, Andrew e Schwab (1994) analisaram a multidimensionalidade do *TES* (*Teacher Effectiveness Survey*), utilizado para medir a efetividade do ensino nos programas de graduação dos professores. Uma amostra de 390 alunos de 33 estados foi usada como fonte de informação. A amostra total foi dividida em duas sub-amostras. A amostra um (1), composta por 195 questionários, foi submetida à análise fatorial exploratória, enquanto a amostra dois (2), composta por 195 questionários, foi utilizada para a validação das hipóteses decorrentes da análise fatorial exploratória através do uso da análise fatorial confirmatória (validade de construto). Os resultados evidenciaram a mesma estrutura fatorial – validade de construto-convergente. As três dimensões altamente correlacionadas foram denominadas como: 1) características didáticas do ensino do professor (instrução); 2) características interpessoal e profissional; e, 3) atualização contínua/liderança. Os autores justificam as intercorrelações entre as dimensões devido à existência de uma única dimensão de ordem superior para a medida da efetividade global do ensino ou pelo fato de que se um professor é excelente num fator, tende também a ser excelente nos demais fatores.

Burdsal e Bardo (1986) analisaram a dimensionalidade do questionário denominado *SPTE* (Percepção dos Estudantes sobre a Efetividade do Ensino), utilizando-se das valorações estudantis de 2115 classes e das técnicas empíricas de análise fatorial ortogonal e oblíqua. Foram encontradas duas soluções: a primeira apresenta uma estrutura fatorial com seis fatores correlacionados de 1ª ordem que explicam 68,8% da variabilidade total do questionário. Esses fatores foram denominados de: 1) atitudes dos estudantes com respeito às características do ensino; 2) carga de trabalho; 3) valor da disciplina para os estudantes; 4) organização e estrutura da disciplina; 5) qualidade da avaliação; e, 6) nível dos materiais; a segunda apresenta uma estrutura fatorial com dois fatores de 2ª ordem,

denominados, respectivamente, de: 1) qualidade geral e 2) dificuldade geral da instrução recebida.

O Questionário de Avaliação de Ensino (*TFQ*) utilizado na cidade universitária de Hong Kong foi analisado empiricamente por Shah *et al.* (2002). A valoração estudantil de 5246 estudantes universitários foi utilizada para a análise da multidimensionalidade do questionário. A análise empírica evidenciou uma estrutura fatorial com quatro fatores consistentes, que explicam 63% da variância total da escala. As dimensões do *TFQ* foram denominadas, respectivamente, como: 1) *performance* do ensino; 2) disciplina desafiante; 3) motivação dos estudantes; e, 4) competência dos estudantes.

Centra (1993), Braskamp e Ory (1994), (*apud* CASHIN, 1995) identificaram seis fatores comumente encontrados nos formulários de avaliação dos alunos:

1. organização/Planejamento da disciplina;
2. clareza/Habilidade de comunicação;
3. interação professor-aluno/relacionamento;
4. dificuldade do curso/carga de trabalho;
5. notas/exames; e
6. auto-avaliação do aprendizado do aluno.

Feldman (1984) apresenta uma lista ampla de 20 dimensões que podem ser utilizadas na elaboração de instrumento, destinado a medir a efetividade do ensino nas instituições de ensino superior. Os itens que compõe as categorias da efetividade do ensino de Feldman são: 1) estimulação do interesse do aluno; 2) entusiasmo do professor; 3) conhecimento da disciplina; 4) bagagem intelectual, 5) preparação e organização da disciplina; 6) clareza expositiva; 7) nível de investigação; 8) sensibilidade para o nível e progresso da classe; 9) clareza dos objetivos e solicitações; 10) valor do material do curso; 11) uso de materiais suplementares; 12) perceber o impacto do aprendizado; 13) imparcialidade;

14) características pessoais; 15) avaliação, 16) discussão; 17) mudança intelectual; 18) respeito para com os alunos; 19) presteza e auxílio; e, 20) dificuldade/trabalho.

Capelleras e Veciana (2000), avaliaram e aplicaram uma escala de medida da qualidade de serviço, no ensino universitário, a partir das percepções de uma amostra de 811 estudantes. As avaliações estudantis foram submetidas à análise fatorial exploratória e confirmatória. Dois modelos foram submetidos à análise fatorial confirmatória: 1º modelo com um fator geral e 2º modelo com cinco fatores correlacionados. Os indicadores de bondade do ajuste apontam que o modelo com cinco dimensões correlacionadas seja o mais adequado para a avaliação da qualidade do ensino e dos serviços no âmbito da universidade. As dimensões validadas do instrumento foram rotuladas como:

Dimensão I – **atitudes e comportamento dos professores**. Essa dimensão está relacionada com a preocupação dos professores quanto ao aprendizado dos alunos, quais sejam: despertar no aluno o interesse pela matéria, estímulo à participação do aluno, disponibilidade para orientação dos trabalhos, utilização de métodos de avaliação adequados e estímulo à comunicação entre alunos e professores.

Dimensão II – **competência dos professores**. Essa dimensão engloba os aspectos relativos ao nível de conhecimentos teóricos e práticos dos professores, atualização dos professores e a sua capacidade para transmiti-los com clareza.

Dimensão III – **integração das disciplinas do curso**. Resume indicadores que fazem referência ao grau de orientação teórico-prática das disciplinas, ao grau em que se fortalece o desenvolvimento das capacidades dos estudantes, à interação entre disciplinas obrigatórias e optativas e à utilidade do material docente.

Dimensão IV – **instalações e equipamentos**. Esse componente tem relação com os elementos tangíveis do serviço, tais como: se as condições das aulas são adequadas para a

docência, se os equipamentos são modernos, se as instalações físicas são cômodas e acolhedoras, e se os livros disponíveis são suficientes.

Dimensão V – **organização do ensino**. Essa dimensão faz referência aos seguintes aspectos: se os horários das aulas são adequados às necessidades dos estudantes e se o número de estudantes por turma é adequado para a docência.

A multidimensionalidade do instrumento de avaliação proposto por Casanueva, Periáñez e Rufino (1997), projetada para avaliar a qualidade do serviço docente prestado pelo professor universitário, foi validada por uma amostra de 358 estudantes de uma mesma disciplina. A submissão dos valores estudantis à técnica de análise fatorial exploratória confirma as cinco dimensões resultantes da análise qualitativa. As dimensões, que explicam 55,7% da variância total da escala, são descritas como: 1) atitudes pessoais do professor; 2) competência expositiva; 3) aspectos objetivos da preparação da aula; 4) avaliação/formalidade; e, 5) demanda de esforço.

Owlia e Aspinwall (1988) apresentaram e avaliaram um instrumento com 19 itens, destinados a medir a qualidade do ensino de Engenharia. A análise empírica da multidimensionalidade do instrumento resulta em quatro dimensões que explicam 60,7% da variabilidade total da escala. As dimensões foram rotuladas como: 1) atitudes do professor; 2) conteúdo; 3) recursos acadêmicos; e, 4) competência.

Marsh e Dunkin (1992), Rowley (1996), (*apud* LI; KAYE, 1998) efetuaram uma revisão de alguns instrumentos destinados à avaliação da efetividade do ensino em sala de aula e concluíram que os instrumentos *SEEQ* (Estudantes Avaliam a Qualidade Educacional; MARSH, 1984), *ENDEAVOR* (FREY, LEONARD e BEATTY, 1975) e Feldman (FELDMAN, 1984) são os mais discutidos na literatura em função de demonstrarem elevada confiabilidade de consistência interna, estabilidade fatorial, capacidade de generalização e indicadores consistentes da validade de construto.

Cohen (1981), em sua meta-análise destinada a estudar a validade das avaliações dos alunos, através de estudos de multiseção, identifica as dimensões abaixo, como as mais comuns nos estudos analíticos realizados.

- 1- **Capacidade:** essa dimensão está associada ao domínio da matéria pelo professor, clareza das explicações e ajuste ao nível da classe.
- 2- **Relacionamento:** esse conceito está associado às seguintes características pelo professor: amigável, flexível, tolerante, está disponível para orientar ou tirar dúvidas dos alunos fora da classe.
- 3- **Estrutura:** este conceito enfatiza a pontualidade do professor e se o professor administra bem o tempo em sala de aula.
- 4- **Dificuldade:** essa dimensão está associada ao nível de dificuldade dos textos.
- 5- **Interação:** esse conceito está associado ao professor que encoraja os estudantes a expressar o seu ponto de vista e facilitador da discussão em classe.
- 6- **Avaliação:** esse dimensão torna claro que o bom professor é aquele que informa aos estudantes quando estes realizam um bom trabalho e informa sobre o seu progresso na disciplina.

Ramos (1997) comparou os instrumentos de avaliação da docência universitária pelos alunos com os estudos desenvolvidos por Trent e Cohen (1973), Marsh (1982 e 1987), Frey, Leonard e Beatty (1975), Ramsden (1991), Tejedor (1993) e Ramos (1996), os quais têm sido suficientemente validados, concluindo que, de uma maneira geral, esses instrumentos de valoração da competência do docente universitário pelos alunos apresentam dimensões que valoram quatro tipos de aspectos:

1. **vinculados ao domínio da disciplina** – Condição necessária, mas não suficiente, para o exercício da docência com qualidade;

2. **didático-técnico.** Esse fator engloba os seguintes indicadores: programação-organização, avaliação - exames, e uso de recursos didáticos metodológicos que favorecem a clareza da exposição da aula e o desenvolvimento da matéria;
3. **comunicação com o aluno.** Fator relacionado com a adequada comunicação do docente tanto em nível grupal como individual e um adequado clima de relação professor-aluno;
4. **motivação dos alunos.** Esse fator está muito relacionado com os fatores um e dois. Inclui a clareza expositiva do professor, o entusiasmo do professor e a sua capacidade de manter os discentes motivados. Esse fator está muito relacionado com a componente vocacional do professor.

As principais conclusões que se podem destacar a partir desses instrumentos de avaliação da docência pelos alunos são as seguintes:

- 1) não existe uma escala padrão generalizada dado que as maiorias dos autores desenvolveram e validaram uma bateria de itens próprios;
- 2) a metodologia utilizada, na elaboração dos instrumentos, parte da análise teórica e lógica; posteriormente a evidência empírica da multidimensionalidade do instrumento é corroborada pela submissão dos dados à técnica de análise fatorial exploratória (componentes principais);
- 3) obtêm-se resultados diferentes quanto ao número de dimensões. O número de dimensões varia entre três e nove. A maioria dos instrumentos apresenta entre quatro e cinco dimensões; e
- 4) de forma sintética, os instrumentos investigados dão ênfase a dois núcleos fundamentais:
 - a) Competência e habilidades dos professores que compreendem as seguintes dimensões: estrutura, organização dos conteúdos, clareza das exposições, habilidades em sala de aula, relevância dos temas.

- b) Atitudes do professor que incluem as seguintes dimensões: relacionamento individual e coletivo, motivação para o aprendizado, carga de trabalho, avaliação justa e exigência.

3 ASPECTOS TEÓRICOS DA CONFIABILIDADE DAS AVALIAÇÕES DOS DOCENTES PELOS DISCENTES

3.1 Introdução

Os problemas em torno do uso dos resultados das avaliações dos docentes pelos discentes, segundo Cashin (1995), estão associados à confiabilidade e a validade das medidas. A confiabilidade da medida é o nível em que a medida está isenta de erros aleatórios. Não obstante, evidências empíricas a favor da confiabilidade da medida não garantem a validade do indicador para medir determinado conceito. Por exemplo, os níveis de satisfação dos alunos, em uma determinada classe, sobre a excelente aparência do professor estão completamente acordados entre eles, oferecendo, assim, uma medida completamente confiável, mas não será válida caso esteja fortemente associado com o conceito de produtividade do professor no contexto da sala de aula. Assim, além da confiabilidade da medida, é necessário que ela expresse um nível de correlação significativo com o conceito que se deseja medir. Portanto, uma medida é confiável e válida quando ela estiver associada com o menor erro possível ao conceito operacionalizado.

Uma escala multidimensional, como a usada pela equipe de Avaliação Institucional e Clima Organizacional da Universidade do Sul de Santa Catarina (UNISUL), deve ter suas propriedades psicométricas de confiabilidade e de validade evidenciados através dos dados empíricos. Indicadores de confiabilidade e de validade estatisticamente significativos evidenciam a alta qualidade técnica do instrumento e, conseqüentemente, maior segurança para os que fazem o uso dos resultados decorrentes do processo avaliativo.

3.2 Erros de medição

Na UNISUL, o aluno universitário é convidado a expressar o seu nível de satisfação sobre um conjunto de variáveis que têm como finalidade medir a qualidade do ensino universitário. No bloco “avaliação de disciplina”, se o aluno estiver completamente satisfeito com a *performance* do professor, em determinado item da escala, ele atribui o grau quatro; se estiver parcialmente satisfeito, ele atribui o grau três, se estiver parcialmente não satisfeito, ele atribui o grau dois e se estiver nada satisfeito, ele atribui o grau um. Como o aluno é convidado a responder 30 itens referentes ao desempenho do docente, determinado professor pode obter 120 pontos, no máximo, caso o aluno esteja completamente satisfeito com o desempenho do professor em todos os itens; ou pode receber 30 pontos, no mínimo, caso o aluno esteja insatisfeito com a *performance* do professor em todos os itens.

Supostamente, esse total de pontos está expressando uma medida que seria a magnitude do desempenho do professor nos diversos itens da escala. Contudo, a medição de qualquer conceito, principalmente o conceito de qualidade de ensino em ambientes educacionais, onde existem muitos fatores que afetam as características da qualidade das medidas, incluindo o processo de medição em si, não está isenta de erros, devido à complexidade do contexto.

O erro de medição é a diferença entre a pontuação observada nos itens do instrumento e a pontuação verdadeira. Caso a medição estivesse totalmente ausente de erro, a pontuação observada de determinado professor seria igual a sua pontuação verdadeira. Assumindo que todos os indicadores valorados pelos alunos apresentem algum grau de erro, a pontuação total atribuída a determinado professor pelos discentes, representa tanto a sua pontuação verdadeira como a pontuação não verdadeira. O que se espera é que a pontuação

total empírica do professor aproxime-se da pontuação total verdadeira à medida que haja um controle das diversas fontes de erro.

Os erros de medida podem ser minimizados pela construção de instrumento de medida bem planejado. Uma solução é optar por medições multivariadas, também conhecidas como escalas somadas, em que diversos itens da escala intercorrelacionados proporcionam uma medida composta do conceito operacionalizado. Segundo Hair *et al.* (1999), o uso de vários indicadores intercorrelacionados proporciona uma medição mais precisa do conceito bem como no estabelecimento dos aspectos teóricos do conceito numa perspectiva mais completa.

3.3 Fontes perturbadoras da confiabilidade

É muito comum, nas apresentações dos resultados das avaliações do curso ou dos docentes, depoimentos de professores da seguinte natureza: o programa de ensino foi entregue e amplamente discutido com todos os alunos da classe e o nível de satisfação recebido nesse item da escala foi inferior a 40%. Outros citam que bastou chegar dez minutos em um único dia do semestre, para que ele recebesse somente 30% de satisfação nesse item da escala. Essas e muitas outras questões se constituem em reclamações justas e pertinentes. Infelizmente, são desconhecidas as razões que levam os alunos a não pontuarem corretamente certos indicadores da escala. Pesquisas adicionais que possibilitem conhecer as causas dessas fontes de perturbações necessitam ser implementadas.

Thorndike (1947, *apud* REQUENA, 1990) apresenta que existem muitas fontes possíveis de onde surgem os erros de medida ou ruído que afetam a confiabilidade das medidas. Classifica as fontes de variabilidades em cinco categorias, quais sejam: características gerais dos indivíduos; características específicas dos indivíduos; características

temporais, porém gerais dos indivíduos; características temporais, porém específicas dos indivíduos; e outras.

Admitindo-se que essas cinco categorias podem ser aplicadas para justificar os erros de medida decorrentes das avaliações dos professores pelos alunos, faz-se a apresentação de tais categorias abaixo:

1. **características gerais dos alunos** - corresponderiam genericamente com as variações entre os níveis de aptidão dos alunos. Nem todos os alunos possuem habilidade geral para compreender a instrução recebida de determinado professor;
 2. **características específicas dos alunos** - faz referência a situações tais como, por exemplo, tendência do aluno em dar respostas constantes para com todos os indicadores da escala;
 3. **características temporais, porém gerais dos alunos** - humor, o estado de saúde, o cansaço, o calor ou ruídos são fatores que, sem dúvida, podem influir sobre as respostas dos alunos. São variáveis temporais sujeitas à variação;
 4. **características temporais, porém específicas dos alunos** - são aquelas circunstâncias temporais que estão relacionadas, como por exemplo, com as flutuações ou confusões da memória dos alunos, forçando a alteração da opinião do aluno sobre alguma variável. Um exemplo é a comum manifestação por parte dos discentes, de que os docentes não entregaram o plano de ensino no primeiro dia de aula, uma vez que, na UNISUL, esse procedimento é norma. “Todo professor deve entregar e discutir o plano de ensino no primeiro dia de aula”; e
1. **outras** - é uma categoria residual, onde se considera que, na maior parte dos casos, se corresponderia com a componente geral de erro não assinalável.

Segundo Moreira (1986), quando o instrumento for particularmente de opinião, em que o aluno evidencia o seu nível de satisfação para com um dado indicador, existem certos mecanismos, talvez desconhecidos, que distorcem as respostas dos alunos.

Ele nos diz que o que acontece, na verdade, é que desconhecemos em grande parte esses mecanismos de adulteração daquilo que talvez pudesse ser chamado de “realidade objetiva”. Esses mecanismos têm maior probabilidade de se fazerem presentes na proporção em que as questões exigem um julgamento do aluno, ainda que sobre fatos teoricamente de seu conhecimento. Por outro lado, a situação complica-se ainda mais se o julgamento depender de critérios de valor do próprio aluno, quando então a noção de “verdade” é influenciada por tais critérios (MOREIRA, 1986, p. 120-121).

3.4 O modelo linear de pontuação

Para compreender melhor todas as fontes de erro que atuam na valoração das indicadoras pelos alunos, pode-se aplicar o modelo de erro de medição, proposto por Spermann (1913, *apud* REQUENA, 1990), o qual constitui a equação fundamental da psicometria e é definido por:

$$X_{ij} = X_i^V + e_{ij} \quad (1)$$

Onde:

X_{ij} representa a *j-ésima* medição observada sobre o *i-ésimo* indivíduo;

X_i^V corresponde à pontuação verdadeira, que nunca será conhecida, sempre estimada do indivíduo *i*; e

e_{ij} é a parcela de erro proveniente da ação das diversas fontes de erro aleatório.

Segundo Requena (1990), na aplicação desse modelo, um conjunto de pressupostos sobre a relação entre a pontuação verdadeira (X_i^V) e o erro aleatório (e_{ij}) devem ser levados em consideração:

a) hipótese de nulidade dos erros. Por esta hipótese, os erros, em média, se anulam. Isso implica que a esperança matemática do erro é zero; ou seja:

$$E(e_{ij}) = 0 \quad (2)$$

b) não existe correlação entre a pontuação verdadeira (X_i^v) e o termo de erro em uma mesma aplicação da escala, ou seja:

$$\rho(X_i^v, \varepsilon_{ij}) = 0 \quad (3)$$

Essa restrição faz-se necessária porque não existe razão para supor que pontuações verdadeiras maiores terão erros positivos, enquanto as pontuações verdadeiras menores terão erros negativos. Se isso ocorre, então existe a presença de erro sistemático, que podem ser controlados. Mas, como se está somente considerando erros aleatórios, então vale a hipótese.

c) não existe correlação entre os erros de distintas indicadoras; ou seja:

$$\rho(\varepsilon_{i1}, \varepsilon_{i2}) = 0 \quad (4)$$

Os erros existentes na valoração do indicador “um”, referente ao indivíduo i , não está relacionado com os erros existentes na valoração do indicador “dois”, referente ao indivíduo i , uma vez que os erros são aleatórios, ou seja, eles independem um do outro. A existência de covariância entre os erros indica a presença de fonte sistemática, que pode ser controlada.

d) não existe correlação entre a pontuação verdadeira (X_i^v) de determinado indivíduo i e o termo de erro em um segundo indicador. Nesse sentido:

$$\rho(X_i^v, \varepsilon_{i2}) = 0 \quad (5)$$

Como no modelo (1), é impossível obter a correlação entre a pontuação observada e a pontuação real ou verdadeira, pois a pontuação real é um valor não diretamente observável, a validação do modelo tem que ser obtido de forma indireta. Para tanto, recorre-se a duas medidas para um mesmo indivíduo i :

$$X_{i1} = \lambda_{i1} X_{i1}^v + e_{i1} \quad (6)$$

$$X_{i2} = \lambda_{i2} X_{i2}^v + e_{i2} \quad (7)$$

Considerando-se todos os pressupostos do modelo (1), e supondo-se que os valores verdadeiros coincidem; ou seja, $X_{i1}^v = X_{i2}^v$, três tipos de medidas podem ser definidos:

- 1) Se $\lambda_{i1} = \lambda_{i2} = 1$ e $\text{Var}(e_{i1}) = \text{Var}(e_{i2})$, diz-se que as duas medidas são paralelas;
- 2) Se $\lambda_{i1} = \lambda_{i2} = 1$ e $\text{Var}(e_{i1}) \neq \text{Var}(e_{i2})$, diz-se que as duas medidas são *tau*-equivalentes; e
- 3) Se $\lambda_{i1} \neq \lambda_{i2}$ e $\text{Var}(e_{i1}) \neq \text{Var}(e_{i2})$, diz-se que as duas medidas são congênicas.

3.5 Medição da confiabilidade

Como visto anteriormente, são diversos os fatores que podem causar erros de medição, ou seja, diferença entre a pontuação verdadeira e a pontuação empírica gerada pelo processo de medição utilizado.

Esses erros de medição são decorrentes de causas aleatórias ou de causas sistemáticas. O erro sistemático afeta a medida de forma constante. Em geral, isso se deve à interferência de fatores estáveis que atuam da mesma forma no escore empírico do professor cada vez que se realiza a medição. Por exemplo, se o docente ministra sua disciplina em uma turma razoavelmente grande de alunos, ele estará sujeito a receber uma valoração baixa no indicador, relacionamento individual, por parte de alguns discentes, uma vez que é provável

que em turmas muito grandes de alunos, o professor tenha dificuldade de se relacionar com todos os alunos da mesma forma. Assim, o tamanho da turma pode enviesar a valoração do professor para baixo nesse item da escala.

O erro aleatório interfere na medida de forma não constante, ou seja, é decorrente da falta de consistência nas respostas dos alunos em medições repetidas. Representam fatores transitórios que afetam a valoração do professor, cada vez que se realiza a avaliação de disciplinas, tais como a fadiga do aluno, ruído externo ou outros fatores.

Portanto, para que o modelo (1) fique bem especificado, o termo de erro sistemático (ϵ_{ij}^s) é acrescentado para possibilitar o entendimento da precisão da medida, e adquire a seguinte forma:

$$X_{ij} = X_i^v + \epsilon_{ij}^s + \epsilon_{ij}^a \quad (8)$$

Em que:

X_{ij} representa a *j-ésima* medição observada sobre o *i-ésimo* indivíduo;

X_i^v corresponde à pontuação verdadeira, que nunca será conhecida, sempre estimada do indivíduo *i*;

ϵ_{ij}^s é a parcela de erro proveniente da ação das diversas fontes de erro sistemático;

e

ϵ_{ij}^a é a parcela de erro proveniente da ação das diversas fontes de erro aleatório.

A distinção entre as fontes de erros sistemático e aleatório que atuam no processo de medição dos indicadores fornece o referencial teórico para distinguir entre a confiabilidade e a validade da medida. Como as diversas fontes aleatórias de erro são as que geram inconstâncias nas medidas em medições repetidas, conseqüentemente, as medidas decorrentes são inconsistentes, logo desfavorecem a confiabilidade da medida. Assim, a confiabilidade da medida decorre da inexistência de fontes aleatórias de erro.

Quanto às fontes sistemáticas de erro, elas interferem no processo de medição dos indicadores de forma constante, logo não produzem inconsistências nas medidas. Como, por exemplo, a presença do professor na sala de aula durante o processo avaliativo. Nesse caso, os alunos com medo de receber uma retalhação por parte do professor acabam atribuindo consistentemente níveis de satisfação maiores que o merecido pelo professor. Assim, existe uma diferença significativa entre a pontuação empírica e a verdadeira pontuação do professor.

Portanto, uma medida terá validade perfeita quando elas forem isentas de erros aleatórios e sistemáticos. Os erros aleatórios dificilmente podem ser controlados durante o processo avaliativo, enquanto que os erros sistemáticos podem ser evitados, como por exemplo, a não permanência do professor na sala de aula durante o processo avaliativo. No caso de turmas grandes de alunos, é necessária a pesquisa adicional para se medir a influência do tamanho da turma nas valorações estudantis e, posteriormente, usar esse coeficiente como um fator de correção para a pontuação dos professores que ministram disciplinas nessas turmas.

O que se espera da medição é que a pontuação empírica X_{ij} seja igual ou muito próxima a da pontuação verdadeira X_j^v , ou seja, que a pontuação total de determinado professor, num determinado item da escala, seja igual à sua pontuação total verdadeira, ou melhor, que as avaliações dos docentes pelos discentes estejam livres de erros sistemáticos e aleatórios. Garantidas essas duas propriedades fundamentais das medidas, quais sejam, a sua confiabilidade e a sua validade, é possível disponibilizar as informações decorrentes do processo avaliativo a todos os professores, para que os mesmos reflitam sobre a qualidade de suas ações, as quais visam especificamente ao progresso do aluno, na disciplina e no curso.

Quando, numa determinada sala, o professor recebe valorações similares por todos os alunos em determinado indicador da escala, diz-se que as valorações apresentam baixa variabilidade interna, indicando que as valorações atribuídas pelos discentes ao docente

são consistentes. À medida que todos os indicadores do instrumento de avaliação apresentam individualmente alta consistência interna e a variabilidade que eles produzem em conjunto for grande, ou seja, quando somadas; então, a escala apresenta níveis de consistência elevado. Quanto mais consistentes forem as valorações atribuídas pelos discentes ao docente em determinado item da escala, menor a presença de fontes aleatórias de erros, resultando em medidas confiáveis. Nesses termos, a confiabilidade de uma medida refere-se ao grau em que a medida está livre de erros aleatórios (PÉREZ; SÁNCHEZ, 2000).

3.6 Medição da confiabilidade – teoria da generalização

Utilizando a Teoria da Generalização, é possível determinar até que ponto pode-se generalizar, com base nas medidas disponíveis dos distintos indicadores do instrumento, e definir a confiabilidade da medida, numa perspectiva prática. Peter (1979, *apud* PÉREZ; SÁNCHEZ, 2000) definem a confiabilidade por um coeficiente ou razão que relaciona a variância das pontuações verdadeiras com a variância das pontuações empíricas, ou seja, aquelas atribuídas pelos alunos ao professor.

Lembrando que a equação fundamental da psicometria denotada no modelo (1) é expressa por:

$$X_{ij} = X_i^V + \varepsilon_{ij} ,$$

$$\text{e que } \text{var}(X_{ij}) = \text{var}(X_i^V + \varepsilon_{ij}) = \text{var}(X_i^V) + 2 \cdot \text{cov}(X_i^V, \varepsilon_{ij}) + \text{var}(\varepsilon_{ij}).$$

Como: $\text{cov}(X_i^V, \varepsilon_{ij}) = 0$ (pressuposto b), conclui-se que a variância das pontuações totais ou empíricas é definida pela expressão:

$$\text{var}(X_{ij}) = \text{var}(X_i^V) + \text{var}(\varepsilon_{ij}) \quad (9)$$

Quando, no modelo (9), a variância do termo erro for nula, resulta que os indicadores foram medidos com total ausência de erros aleatórios, resultando em medidas completamente confiáveis. Segundo Peter (1979, *apud* PÉREZ; SÁNCHEZ, 2000), a confiabilidade da medida resulta da comparação da variância das medidas verdadeiras com a variância das medidas empíricas. Portanto, pode-se obter a confiabilidade da medida por:

$$\rho_{X_{ij}} = \frac{\text{var}(X_i^V)}{\text{var}(X_{ij})} \quad (10)$$

Como a variância da pontuação verdadeira não pode ser obtida diretamente e que

$\text{Var}(X_i^V) = \text{var}(X_{ij}) - \text{var}(\epsilon_{ij})$, a equação (10) pode ser reescrita como:

$$\rho_{X_{ij}} = \frac{\text{var}(X_{ij}) - \text{var}(\epsilon_{ij})}{\text{var}(X_{ij})} = \frac{\text{var}(X_{ij})}{\text{var}(X_{ij})} - \frac{\text{var}(\epsilon_{ij})}{\text{var}(X_{ij})}$$

assim,

$$\rho_{X_{ij}} = 1 - \frac{\text{var}(\epsilon_{ij})}{\text{var}(X_{ij})} \quad (11)$$

A equação (12) estabelece que a confiabilidade da medida depende exclusivamente da quantidade de erro aleatório presente na obtenção da medida. Quanto menor a variância do erro, maior a confiabilidade da medida. Se $\text{var}(\epsilon_{ij})$ for nula, a confiabilidade é igual a um, quando a $\text{var}(\epsilon_{ij}) = \text{var}(X_{ij})$, a confiabilidade é nula.

Portanto, essa medida varia de zero a unidade, isto é: $0 \leq \rho_{X_{ij}} \leq 1$. Nesse caso, o procedimento analítico de análise de variância pode ser utilizado para as estimativas da variância do termo de erro (ϵ_{ij}) e da pontuação total (X_{ij}) para o diagnóstico da confiabilidade das medidas.

3.7 Medição da confiabilidade – o coeficiente α de Cronbach

O coeficiente α de Cronbach é o método mais utilizado para medir a confiabilidade, quando a confiabilidade for entendida como uma consistência interna dos indicadores da escala; ou seja, os indicadores da escala, altamente intercorrelacionados, devem estar medindo o mesmo construto latente. Particularmente adequado para instrumentos multidimensionais e segundo Moreira (1986), amplamente utilizado para a estimativa da confiabilidade dos instrumentos de avaliação de professores.

O coeficiente α , proposto por Cronbach (1951), é uma generalização do coeficiente α de Kuder-Richardson, proposto para variáveis dicotômicas. Considera a homogeneidade dos itens da escala e apresenta como vantagem o fato de precisar de uma única aplicação do instrumento avaliativo a uma amostra representativa de alunos. Sua formulação é a seguinte:

$$\hat{\alpha} = \frac{n}{n - 1} \left(1 - \frac{\sum \hat{\sigma}_i^2}{\hat{\sigma}_t^2} \right) \quad (12)$$

em que

n : número de indicadores associado ao construto;

$\sum \hat{\sigma}_i^2$: soma das variâncias dos n indicadores associados ao construto; e

$\hat{\sigma}_t^2$: variância da pontuação total dos n indicadores no construto.

A equação (12) possibilita concluir que o índice de confiabilidade das medidas depende das variâncias individuais dos indicadores e da variância que elas reproduzem no total. Assim, quanto menor forem os valores das variâncias individuais e, quanto maior for o total da variância que os indicadores reproduzem em conjunto, maior é a confiabilidade das medidas. Note que à medida que reduz o somatório das variâncias individuais dos

indicadores, aumenta a variância que esses indicadores possuem em comum, garantindo, assim, a consistência interna do construto.

O α de Cronbach assume valores entre “zero” e “um”. Será igual a “um” quando os indicadores do construto, individualmente, não manifestarem variância. Será igual a “zero” quando o somatório das variâncias dos indicadores for igual à variância do total das pontuações desses indicadores no construto.

Segundo Pereira (1999), o coeficiente α de Cronbach pode ser entendido como o coeficiente de correlação ao quadrado (R^2) entre o construto e seus indicadores, e representa uma suposta medida real do construto em análise. Por exemplo, para um α de 0,95, sugere que 95%, da variabilidade total do construto teórico é explicado pela influência dos indicadores associados ao construto.

Destaca-se que o coeficiente α de Cronbach depende da quantidade de indicadores associados ao construto. Isso significa que o pesquisador pode aumentar a confiabilidade da medida simplesmente aumentando o número de indicadores do construto. Para Hair *et al.* (1999), os pesquisadores não devem aumentar a confiabilidade do construto pela inclusão de mais indicadores. Em se tratando de instrumentos multidimensionais, sugere-se que, para efeitos práticos, os construtos devam ser representados por um número entre cinco e sete indicadores altamente intercorrelacionados e que o fator latente responsável por esses intercorrelacionamentos exceda a cota mínima de consistência interna estabelecida por Nunnally (1987), ou seja, superior a 0,7.

Quando o instrumento de avaliação da docência pelos alunos é destinado a medição de um conceito que possui mais de uma dimensão, a confiabilidade total da escala pode ser obtida pelo coeficiente de confiabilidade β de Raju (1977), que é uma generalização do coeficiente α de Cronbach. Sua formulação é:


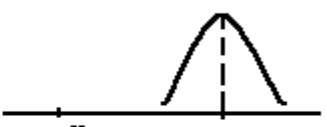
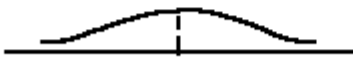
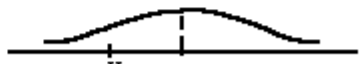
$$\hat{\beta} = \frac{\sigma_t^2 - \sum_{j=1}^K \sigma_j^2}{\sigma_t^2 \left[1 - \sum_{j=1}^K \left(\frac{n_j}{n} \right)^2 \right]} \quad (13)$$

Na equação (13), σ_t^2 é a variância estimada total da escala geral; σ_j^2 corresponde à variância de cada dimensão; n_j é o número de indicadores de cada dimensão; n corresponde ao total de indicadores da escala; e K é o número de dimensões da escala. Essa medida, também, assume valores entre zero e um. Valor superior a 0,7 indica que se trata de uma escala consistente.

3.8 Relação entre validade e confiabilidade

Confiabilidade e validade são duas propriedades das medidas que estão muito estreitamente relacionadas, desempenhando papéis complementares. A confiabilidade de consistência interna está relacionada com a homogeneidade das respostas dos distintos avaliadores (alunos), em uma mesma classe, enquanto que a validade está associada ao grau de certeza que se tem sobre o conceito medido pelos discentes.

Uma medida pode apresentar elevado grau de confiabilidade de consistência interna e estar medindo um conceito não operacionalizado no instrumento. O quadro 1, adaptada de Pérez e Sánchez (1999), apresentada abaixo, relaciona as quatro situações relacionadas com a influência dos erros sistemáticos e aleatórios, respectivamente, sobre a validade e a confiabilidade das medidas. Os efeitos dessas fontes de erro sobre as medidas podem ser percebidos, observando a comparação entre a pontuação verdadeira (X_i^V) e a pontuação empírica atribuída pelos alunos a determinado professor (X_{ij}).

Erro Aleatório (ε_{ij}^a)	Erro Sistemático (ε_{ij}^s)	
	baixo	alto
baixo	 <p>$X_i^v = X_{ij}$ a) válido e confiável</p>	 <p>X_i^v X_{ij} b) confiável e não válido</p>
alto	 <p>$X_i^v = X_{ij}$ c) não confiável, mas válido</p>	 <p>X_i^v X_{ij} d) não confiável, não válido</p>

Quadro 1: Comparação entre confiabilidade e validade, segundo o tipo de erro

A primeira casela (a), reportada no quadro 1, corresponde à situação ideal.

Representa a qualidade de um instrumento com elevada capacidade técnica para a obtenção de medidas confiáveis e válidas. Esse fato pode ser constatado ao observar a pontuação média real de determinado professor com aquela atribuída pelo aluno; observa-se que elas são iguais. A casela (b) mostra um instrumento de medida com alta confiabilidade, pois evidencia alta consistência interna em torno de (X_{ij}), mas acusa a presença de variáveis que induzam ao erro sistemático dado que a distância entre X_i^v e X_{ij} é razoavelmente grande. Na casela (c), a pontuação empírica (X_{ij}) e a pontuação verdadeira (X_i^v) são iguais, mas a presença de erros aleatórios atribui baixa consistência interna para as medidas. Nesse caso, as medidas decorrentes da valoração do instrumento não são confiáveis, para garantir a validade das medidas. A casela (d) é o caso em que o grau de confiabilidade e de validade da medida é baixa, devido à presença de erros aleatórios e sistemáticos, atuando no processo de medição de forma acentuada.

3.9 Pesquisas sobre a confiabilidade da avaliação dos docentes pelos discentes

Na literatura de medição educacional, a confiabilidade está relacionada com a consistência interna, estabilidade e generalização das avaliações dos professores pelos alunos (CASHIN, 1988, 1995).

A confiabilidade de consistência interna, calculada pelo coeficiente α de Cronbach (1951), é um índice que indica o quanto os alunos consistentemente avaliam o professor quando o instrumento é aplicado somente uma vez, numa determinada classe. Esse índice varia entre zero e um. Um valor zero ou próximo indica que não há consistência nas respostas dos alunos, numa determinada classe. Valor um ou muito próximo indica quão consistentemente são as respostas dos alunos numa determinada classe. Segundo Cashin (1988, 1995), a confiabilidade de consistência interna é decorrente do número de avaliadores da classe. Quanto mais avaliadores, maior é o índice de consistência interna. Segundo Nunnally (1987), o índice de consistência interna deve superar 0,7. Indicadores de confiabilidade abaixo de 0,7 devem ser interpretados com precaução (CASHIN, 1995).

A estabilidade está relacionada com a concordância dos índices de correlação das avaliações dos docentes pelos discentes quando os mesmos alunos avaliam o mesmo professor, em ocasiões distintas. A generalização das avaliações dos docentes pelos discentes indica até que ponto as conclusões decorrentes da avaliação, baseada em uma amostra de alunos, aplica-se à população de alunos como um todo.

.Sixbury e Cashin (1995^a *apud* CASHIN, 1995), informam que as confiabilidades médias para os 38 itens do instrumento IDEA (Sistema de Determinação do Progresso do Ensino pelos alunos) são como as apresentadas na tabela 1, apresentada a seguir:

**Tabela 1: Universidade do estado de Kansas
Indicadores de Confiabilidade do Instrumento
IDEA.**

Número de avaliadores	Confiabilidade
10	0,69
15	0,83
20	0,83
30	0,88
40	0,91

Fonte: Cashin, 1995.

A tabela 1, evidencia que, quanto maior for a quantidade de avaliadores numa determinada classe de alunos, maior é o índice de confiabilidade da consistência interna.

Os indicadores da confiabilidade da consistência interna para as dimensões do instrumento de Marsh (1984), “Avaliações do Aluno sobre a Qualidade Educacional” (*SEEQ*), estão apresentadas na tabela 2:

**Tabela 2: Universidade de Sydney – Austrália
Indicadores da confiabilidade do Instrumento SEEQ.**

Números de avaliadores	Confiabilidade
2	0,23
5	0,6
10	0,74
25	0,90
50	0,95

Fonte: Marsh, 1984.

A comparação das tabelas 1 e 2 evidencia que dois instrumentos, mesmo tendo concepções diferentes, valorados por classes com quantidades de alunos similares, resultam em confiabilidades também similares. Segundo Cashin (1995), confiabilidades altas são obtidas com instrumentos bem projetados por especialistas em medição educacional.

Os instrumentos de avaliação da docência universitária, utilizado por Silveira *et al.* (1984), Tejedor *et al.* (1988), Sánchez (1988), Vanleeuwen *et al.* (1999), Capelleras e Veciana (2000), Shah *et al.* (2002) e Medina (2002), relatam elevados índices de confiabilidade de consistência interna de seus instrumentos e seguem apresentados na tabela 3.

Tabela 3: Instrumentos de Avaliação da docência universitária. Indicadores da confiabilidade.

Autor	Confiabilidade (α de Cronbach)
Silveira <i>et al.</i> (1984)	0,98
Tejedor <i>et al.</i> (1988)	0,98
Sánchez (1988)	0,97
Vanleeuwen (1999)	0,97
Capelleras e Veciana (2000)	0,87
Shah <i>et al.</i> (2002)	0,93
Medina (2002)	0,94

Fonte: Marsh, 1984.

Os coeficientes de confiabilidade, reportados nas tabelas 1, 2 e 3, mostram razoável acordo entre todos os alunos diferentes dentro de uma mesma classe, indicando que as medidas usadas por esses pesquisadores apresentam erros aleatórios minimizados.

Silveira, Moreira e Nunes (1985) acrescentam dois itens ao instrumento de Silveira e Moreira (1984) e o submeteram a 1982 alunos que avaliaram o desempenho didático de 36 professores do Instituto de Física da PUCRS. Os coeficientes alfa, obtidos para cada professor, são todos elevados e maiores do que 0,8. A análise de variância com os escores totais possibilitou concluir que o instrumento possui a capacidade de discriminar tipologias de professores.

Um fator ou dimensão é considerado consistente quando todos os seus indicadores integrantes são intercorrelacionados entre si, no esforço de medir determinado conceito teórico. A confiabilidade dos fatores do instrumento (*SEEQ*) de Marsh (1982) e do instrumento (*ENDEAVOR*) de Frey, Leonard e Beatty (1975) foi investigadas por Marsh (1984) em universidades espanholas e australianas. A tabela 4, apresenta os indicadores de confiabilidade das dimensões destes dois instrumentos.

Tabela 4: Confiabilidade dos fatores dos instrumentos *SEEQ* e *ENDEAVOR* – Universidades: Austrália e Espanha.

Dimensões	Austrália	Espanha
<i>SEEQ</i>		
1. Interação em grupo	0,94	0,94
2. Aprendizado/valor	0,92	0,92
3. Trabalho/dificuldade	0,91	0,79
4. Exames/grau	0,81	0,85
5. Relacionamento individual	0,93	0,90
6. Organização/clareza	0,93	0,91
7. Entusiasmo	0,95	0,92
8. Aprofundamento dos temas	0,88	0,89
9. Tarefas/leituras	0,84	0,84
<i>ENDEAVOR</i>		
1. Discussão em classe	0,85	0,92
2. Conquistas dos alunos	0,85	0,87
3. Carga de trabalho	0,94	0,91
4. Avaliação	0,90	0,94
5. Atenção pessoal	0,90	0,91
6. Clareza da aula	0,92	0,89
7. Organização/planejamento	0,85	0,78

Fonte: Marsh, 1984.

Os resultados expostos, na tabela 4 indicam que todas as dimensões do instrumento *SEEQ* e do instrumento *ENDEAVOR*, aplicados em contextos distintos, medem consistentemente os conceitos para os quais foram designados para medir. Os elevados índices de consistência interna desses dois instrumentos fornecem o suporte empírico para a estabilidade e a generalização das escalas.

Um outro instrumento que tem sido extensivamente analisado é o Questionário de Experiência do Curso (*CEQ*) de Ramsden (1991). As confiabilidades dos fatores do *CEQ* estão apresentadas segundo a dimensão, na tabela 5.

Tabela 5: Confiabilidade dos fatores do *CEQ* Universidade de Melbourne

Dimensões	Confiabilidade
1. Ensino adequado	0,87
2. Clareza dos objetivos do curso	0,80
3. Adequada carga de trabalho	0,77
4. Avaliação adequada	0,71
5. Estímulo/motivação dos alunos	0,72

Fonte: Ramsden (1991).

Os indicadores da consistência interna da escala de Ramsden (1991), apresentados na tabela 5, indicam que a escala tem uma boa qualidade estatística, uma vez que as consistências internas de todas as dimensões do *CEQ* superaram a 0,7. Esses coeficientes de confiabilidade indicam que os alunos, consistentemente, reconhecem as cinco dimensões designadas para medir a *performance* dos cursos de graduação.

O estudo da dimensionalidade, por técnicas de análise fatorial exploratória e confirmatória, do instrumento de avaliação da qualidade dos serviços universitários de Capelleras e Veciana (2000), confirmou cinco dimensões que apresentam elevados índices de consistência interna e seguem apresentadas na tabela 6. Esses indicadores manifestam que

todas as dimensões da escala não apresentam problemas quanto à presença de fontes aleatórias de erros que atuam durante o processo de obtenção das medidas.

Tabela 6: Confiabilidade das dimensões da escala.

Dimensões	Alpha de Cronbach	Confiabilidade Composta
I	0,8515	0,8544
II	0,8506	0,8540
III	0,8347	0,8423
IV	0,7775	0,7808
V	0,7433	0,7601

Fonte: Capelleras e Veciana, 2000.

Os indicadores de consistência interna, reportados nas tabelas 4, 5 e 6, respectivamente, revelam quão consistentemente são as respostas dos alunos, indicando que se trata de instrumentos confiáveis. A confiabilidade é uma condição indispensável a qualquer instrumento de avaliação, dado que não é possível testar hipóteses a partir de medidas inconstantes.

A confiabilidade composta do construto é o indicador da consistência interna das dimensões, testadas pelas técnicas de análise fatorial confirmatória. O limite aceito para essa medida de confiabilidade composta é de 0,7 (HAIR *et al.* 1999).

3.10 Estabilidade das avaliações

Conforme Aubrecht (1981), coeficientes de estabilidade são medidas de concordância das avaliações dos mesmos alunos, para com o mesmo professor, com o tempo. Segundo Sánchez (1985), as correlações significativas entre as avaliações dos alunos obtidas

na metade e no final da disciplina, ou entre avaliações de ex-alunos e avaliações de alunos atuais, indicam a estabilidade da avaliação do professor.

Costin (1968, *apud* COSTIN *et al.* 1971) estima correlações significativas entre os níveis de satisfação dos alunos com relação à qualidade do ensino, recebido na metade do semestre e no fim do semestre. A correlação mínima e máxima para as cinco dimensões da escala quando comparadas foram, respectivamente, iguais a 0,48 e 0,87. Esses resultados evidenciam a estabilidade das avaliações estudantis. Como foi obtida congruência dos resultados, tem-se mais evidências de quão confiante são as valorações estudantis sobre os seus professores.

Overall e Marsh (1980) estudaram os níveis de satisfação dos alunos no final de 100 disciplinas e mais uma vez, pelo menos, um ano após os mesmos alunos terem se graduados. O coeficiente de estabilidade médio encontrado foi de 0,83. Greenough e MENGES (1971), Gillmore (1973), Hogan (1973), (*apud* TEJEDOR, 1988) mostram que os níveis de satisfação dos alunos com respeito a um mesmo professor correlacionam com valores que vão desde 0,70 a 0,87.

Firth (1979, *apud* SANCHEZ, 1985) solicitou aos estudantes já graduados e que estivessem durante um ano trabalhando, que valorassem a efetividade de seus professores, novamente, e que identificassem os melhores e piores professores. A diferença entre ambas as valorações intertemporais não foi considerada significativa.

As correlações dos trabalhos de Costin (1978), Marsh e Overall (1980) e Firth (1979) fornecem evidência empírica para a estabilidade das avaliações dos alunos e possibilitam sustentar a hipótese de que os alunos fazem julgamentos consistentes sobre o desempenho do professor.

Marsh e Hocevar (1991) estudaram a estabilidade das nove dimensões do SEEQ (Estudantes Avaliam a Qualidade Educacional), trabalhando com os níveis de satisfação dos

alunos de 24158 disciplinas distribuídas em 21 subgrupos. Os subgrupos foram construídos considerando o nível de instrutor, nível do curso e disciplina acadêmica. As mesmas nove dimensões bem definidas e consistentes do *SEEQ* foram encontradas para as 22 análises fatorial realizadas, estratificada segundo o nível do instrutor, nível do curso e disciplina acadêmica. Os resultados fornecem evidências empíricas para a estabilidade e capacidade de generalização para as nove dimensões do *SEEQ*.

Estudos em universidades brasileiras referentes à estabilidade dos fatores do formulário de avaliação dos docentes e referentes à estabilidade das avaliações dos docentes foram feitos por Moreira (1986). Neste estudo², foram levantadas e comparadas evidências empíricas para a estabilidade das dimensões fatoriais do ensino, analisando as avaliações dos professores pelos alunos ocorridos nos anos de 1982 e 1985. Os resultados mostram que os fatores retidos pela análise fatorial exploratória são apenas parcialmente semelhantes. Quanto à estabilidade da avaliação do professor, o estudo concluiu que não há, a priori, uma constância na avaliação. Enquanto alguns professores permanecem estáveis nas suas avaliações, outros diminuíram e outros aumentaram.

3.11 Generalização das avaliações

O estudo da generalização diz respeito ao grau de confiança com que os dados resultantes dos processos de avaliação reflitam a qualidade do ensino do professor em qualquer semestre letivo. Decisões pessoais implicam em previsões sobre a qualidade do ensino daquele professor em cursos futuros. Por isso, o termo generalização, um conceito que incorpora os conceitos de confiabilidade e validade (AUBRECHT, 1981).

² Este é o único estudo publicado no Brasil até a data desta defesa, que trata da estabilidade da estrutura fatorial de instrumento de avaliação da docência pelos alunos e da estabilidade das avaliações dos alunos com respeito ao corpo docente ao longo do tempo.

Nos estudos de generalização, duas abordagens têm sido empregadas pelos pesquisadores: a Teoria da Confiabilidade Clássica e a Teoria da Generalização.

A Teoria da Confiabilidade é usada para analisar as correlações nos níveis de satisfação dos alunos sob as seguintes condições:

- 1- mesma disciplina – mesmo professor (MD – MP);
- 2- mesma disciplina – diferente professor (MD – DP); e
- 3- diferente disciplina – mesmo professor (DD – MP).

A condição MD-MP controla o efeito classe que é uma combinação do efeito professor, efeito disciplina e a interação entre disciplina e professor. A condição MD-DP serve para isolar o efeito disciplina, enquanto a condição DD-MP serve para isolar o efeito professor.

A Teoria da Generalização usa a técnica de análise de variância para examinar a quantidade de variância nos níveis de satisfação dos alunos devido ao “efeito professor”, “efeito disciplina”, “efeito classe” (AUBRECHT, 1981).

Hogan (1973), Bausell *et al.* (1975), Marsh (1982), (*apud* AUBRECHT, 1981) usaram a teoria de confiabilidade para estudar os efeitos diferenciais do professor e da disciplina, correlacionando os níveis de satisfação dos alunos nas três categorias diferentes MD-MP, MD-DP, e DD-MP. As correlações encontradas por Hogan (1973) e Bausell *et al.* (1975) e Marsh (1982) seguem apresentadas na tabela 7. O valor entre parênteses corresponde ao percentual da variância explicada para cada uma das combinações.

Tabela 7: Correlações do Estudo de Hogan (1973), Bausell *et al.* (1975) e Marsh (1982).

Correlações Efeitos	Hogan	Bausell <i>et al.</i>	Marsh
DD – MP (professor)	0,40 (16%)	0,37 (13,69%)	0,52 (27,04%)
MD–DP (disciplina)	0,19 (3,61%)	0,24 (5,76%)	0,14 (1,96%)
MD–MP (classe)	0,70 (49%)	0,64 (40,96%)	0,71 (50,41%)

Fonte: Aubrecht (1981).

Os resultados das pesquisas de Marsh (1984), Hogan (1973), Bausell *et al.* (1975), Gillmore *et al.* (1978) apontam evidências de que o professor, não a disciplina, fornece argumentos para os níveis de satisfação dos alunos; possibilitando, então, a prática de comparar avaliações de professores diferentes para decisões diversas (aperfeiçoamento, estabilidade, promoção, etc.).

4 VALIDADE DAS AVALIAÇÕES DOS DOCENTES PELOS DISCENTES

4.1 Introdução

Para Churchill (1979), uma medida tem validade quando as diferenças entre os valores observados refletirem somente diferenças verdadeiras sobre as características que se pretende medir e não sobre outros fatores. Por exemplo, dois professores A e B obtiveram respectivamente a pontuação nove e sete quanto à sua avaliação no construto “avaliação”. Se a medida for válida, a diferença de dois pontos percentuais observada, deve-se aos atributos da boa docência que distingue um professor do outro. O professor “A”, por exemplo, é aquele que elabora a avaliação numa linguagem clara, fixa os critérios de correção, elabora a avaliação com vistas aos objetivos acordados com a turma; enquanto que o professor “B” manifesta essas qualidades, mas em menor intensidade. No entanto, se a pontuação nove obtida pelo professor é decorrente de uma troca, ou seja, o professor consegue iludir os alunos com notas altas para obter melhores níveis de satisfação, há uma situação de fraqueza para os resultados da avaliação, resultando em informações não válidas com vistas à melhora da qualidade do ensino.

É necessário o estudo da confiabilidade e da validade dos itens que compõem a escala e do instrumento como um todo, para evidenciar a qualidade das medidas decorrentes do processo avaliativo, ou seja, comprovação de que as medidas estão isentas das diversas fontes de erro que atuam durante o processo de medição. Peter (1981), considera que a medida de um conceito é válida quando cumpre os seguintes requisitos: a) medirem a grandeza e a direção de uma amostra representativa de indicadores do conceito e b) a medida não está enviesada com elementos procedentes do domínio de outro construto ou com erros.

Para Ramos (1986), o conceito de validade é fundamental na investigação científica, pois dessa característica primordial da medida decorrem os conceitos de contraste, verdade, adequação à realidade, utilidade dos resultados, que estão todos vinculados ao conceito de validade da medida. Segundo Magnuson (1966, *apud* RAMOS, 1986), um teste ou instrumento de medição cuja confiabilidade é conhecida, será válido se medir os traços para o qual foi construído. Por essa definição, fica claro que, para estimar a validade do instrumento de medida, é necessário o conhecimento prévio dos conceitos que se espera que o instrumento seja capaz de medir. Assim, se existir, no instrumento, um conjunto de variáveis ou itens intercorrelacionados e essas convergirem, segundo os procedimentos empíricos de análises de dados, para a direção do conceito investigado, diz-se que esse conjunto de indicadores é válido para auxiliar o pesquisador no estabelecimento pleno do conceito. Em sentido restrito, um instrumento de medida é válido se cumpre satisfatoriamente ao propósito para o qual foi construído, nesse caso, avaliar a qualidade do ensino no âmbito da sala de aula.

Para Ramos (1986), o tema validade é complexo e justifica a sua afirmação pelas seguintes razões:

- 1) o caráter indireto da medição, já que nunca se mede uma variável latente. O que se mede são algumas manifestações do domínio de condutas, que se supõem serem manifestadas pelos indivíduos. Nesse sentido, fica difícil manifestar se o instrumento é válido ou não válido, e sim em que grau é válido para medir o que se pretende medir, dado que somente uma amostra de itens é considerada para medir determinado traço latente; e
- 2) o que possibilita a determinação do grau de validade do instrumento de medida são as avaliações atribuídas aos diversos itens do instrumento por uma amostra representativa de respondentes. Assim, a estimativa do grau de validade do instrumento passa a ser uma característica da população considerada. Desse modo, um instrumento pode ter um certo

grau de validade para uma população de indivíduos e não ser válido ou não apresentar o mesmo grau de validade para outra população de respondentes distinta.

4.2 Tipos de validez

A validez não é um conceito único, adota-se diversas e variadas formas de interpretações. Depende das variáveis a validar, dos objetivos do instrumento de medida, da população a ser submetida ao instrumento e outras. Considerando que as medidas coletadas podem ter diferentes finalidades, existem também diferentes tipos de validez, quais sejam: de conteúdo, de construto, convergente, discriminante, fatorial, preditiva e concorrente. Segundo Tourón (1989) os tipos de validez foram amplamente discutidos pela Associação Americana de Psicologia em 1954, 1966, 1974, 1985 e 1986.

A primeira versão é a de 1954, que fixa os termos de validez de conteúdo, de critério (concorrente e preditiva) e de construto.

A validez de conteúdo refere-se à relevância e representatividade de um conjunto de variáveis de um instrumento destinado a medir determinado conceito previamente definido. As pesquisas que tratam da validez de conteúdo preocupam-se em responder a seguinte pergunta: em que grau essa escala apresenta uma amostra representativa de variáveis ou indicadores relacionados ao conceito a ser medido?

A validez de critério está relacionada com o nível de correlação existente entre as distintas dimensões do instrumento com outras variáveis relacionadas como critérios significativos da qualidade do ensino. Conforme o prazo envolvido, a validade de critério pode adquirir duas formas distintas: validade concorrente e preditiva. Verifica-se a validade concorrente, quando os dados sobre a escala que está sendo avaliada e sobre as variáveis de critério são coletadas conjuntamente. Por exemplo, pode-se estimar o grau de validade

concorrente entre a dimensão “exigência” do instrumento de avaliação da docência com a variável critério “o seu curso está exigindo o suficiente para a sua formação profissional”. Quando essas informações são coletadas simultaneamente e se existir entre elas um índice de correlação positivo e significativo, tem-se o suporte empírico para a validade concorrente.

A validade preditiva refere-se à capacidade da escala para prever o valor futuro de outras variáveis. Para avaliar a validade preditiva, recomenda-se a coleta de dados sobre a escala, em um determinado momento, e dados sobre as variáveis de critério em um tempo futuro. Por exemplo, estimar o grau de validade preditiva entre a dimensão “exigência” do instrumento de avaliação da docência e as médias finais dos alunos, de determinada classe. Dado que as informações são coletadas em momentos diferentes e existir entre essas variáveis, correlação positiva e significativa, tem-se o suporte empírico para a validade preditiva do instrumento ou da escala.

Cronbach e Meehl (1967, *apud* RAMOS, 1986), cita que esses autores definem um construto como algum atributo postulado sobre as características das pessoas, que supostamente sejam confirmadas ou não pelos respondentes na aplicação do instrumento. Em se tratando do formulário de avaliação da docência, um construto é alguma característica da qualidade dos docentes que se refletirá nos resultados provenientes de uma aplicação do instrumento. Construto é o termo usado para definir um conceito que não pode ser diretamente e perfeitamente medido. O que se mede são os indicadores, agentes de construto. Tinsley e Tinsley (1987), escreveram que fatores são “construtos ou teorias hipotéticas que ajudam a interpretar a consistência num conjunto de dados”. Kline (1994), definiu um fator como uma dimensão ou construto que é um conceito teórico resultante do relacionamento de um conjunto de variáveis. Kim e Meuller (1978) citam que os fatores são variáveis subjacentes, hipotéticas e não medidas diretamente, que se presume serem as fontes dos

indicadores. O número de construto de um dado instrumento de avaliação são bem menores do que o número de indicadores e são responsáveis pela covariância entre os indicadores.

A validade do construto depende da importante relação entre o conceito teórico e os indicadores. Por exemplo, se os alunos reconhecem através dos seus níveis de satisfação que os indicadores: i) o professor redige provas ou verificações da aprendizagem de forma clara; ii) o professor informa sobre os critérios adotados na avaliação; e, iii) o professor planeja as avaliações compatíveis com os objetivos e conteúdos ministrados estão intercorrelacionados, de modo a constituir consistentemente um construto ou dimensão latente, então a dimensão “avaliação” operacionalmente definida está bem evidenciada.

A validez de construto se aplica porque o instrumento não é um conjunto de indicadores inconexos, mas um instrumento projetado para medir algo concreto. O indicativo de validade de construto torna-se necessário quando o usuário do instrumento deseja realizar inferências a partir das pontuações do instrumento em relação às indicadores que podem ser agrupadas sob um rótulo de um construto teórico em particular ou para determinar a quantidade de construtos teóricos que explicam os resultados de um instrumento de avaliação.

A validez de construto refere-se a um teste, quando é interpretado como uma medida de algum atributo ou qualidade que não está operacionalmente definido. O problema consiste em dar resposta a seguinte pergunta: quais construtos são responsáveis pela máxima variabilidade dos resultados de um teste?

Segundo Tourón (1989), a segunda versão das recomendações da Associação Americana de Psicologia (APA) é de 1966, a qual amplia o conceito de validez de construto, deixando claro que quando se está validando um construto, está validando todo um conjunto de teorias nas quais o construto está inserido, de modo que é toda uma rede de relações que será submetida à prova.

A terceira revisão feita pela Associação Americana de Psicologia (1974), trata dos padrões de elaborações de testes. Esclarece que validade não é uma característica do instrumento de medida e sim das interpretações e das inferências que se fazem com as pontuações obtidas a partir da valoração do instrumento. Reporta que a validade se reduz a dois tipos básicos: a validade relacionada com o significado ou natureza do conceito medido – validade de construto e a validade relacionada com o uso do instrumento como indicador de outras variáveis ou validade preditiva. Atualmente, a validade de construto engloba todos os tipos de validade, sendo a validade preditiva uma de suas partes. A validade preditiva ajuda a estabelecer as relações entre os construtos e outras variáveis critério ou construtos (TOURÓN, 1989; HAIR *et al.* 1999; ANASTARI, 1986).

A quarta versão, a de 1985, fixa o processo operativo para o estabelecimento de validade de construto mediante a análise de validade convergente e da validade discriminante através de métodos analíticos (TOURÓN, 1989).

Para Pérez e Sánchez (1999), validade convergente é o tipo de validade de construto que avalia o grau em uma escala, se correlaciona significativamente com outras medidas projetadas para medir o mesmo conceito. Ainda, segundo o mesmo autor, validade discriminante é um tipo de validade de construto que avalia o grau em uma escala, não se correlaciona com outra, segundo as teorias da natureza do construto. Os procedimentos básicos para comprovar esses tipos de validade são as matrizes de multimétodos-multitraços e pela análise fatorial confirmatória (PÉREZ; SÁNCHEZ, 1999).

Segundo Ramos (1986), são os construtos que determinam que indicadores se encontram relacionados para sua observação. A validade de construto é que indicará o grau em que o instrumento de avaliação é uma medida adequada do construto e em que medida as hipóteses derivadas do construto podem confirmar-se mediante a utilização do instrumento. Atualmente, por ser o caminho mais integrador dos diferentes tipos de validade, a validade de

construto, com suas múltiplas técnicas e procedimentos, é a trajetória adequada para a validação do instrumento de medida. Tem como propósito fundamental validar a teoria subjacente ao sistema de avaliação ou de medida, seja esta uma teoria da personalidade ou de conduta dos indivíduos (CRONBACH; MEEHL, 1967, *apud* RAMOS, 1986).

Em se tratando de questionário de avaliação da docência pelos alunos, que engloba um conjunto de características que, *a priori*, definem um bom professor, características estas que não são redutíveis uns aos outros, torna-o multidimensional. Portanto, o tipo de validade a ser evidenciado é o de validade de construto, ou seja, analisar através dos dados procedentes dos alunos se esses reconhecem as distintas dimensões do instrumento de avaliação.

O quadro2, adaptado de Pérez e Sánchez (1999), apresenta os tipos de validade requerida para com os instrumentos de avaliação da docência, com suas características básicas e métodos mais usuais para quantificar esse tipo de validade.

Validez	Características básicas	Métodos
De conteúdo	É utilizado para garantir uma amostra representativa de indicadores consistentemente relacionados com o conceito a ser medido.	Correlação item-total.
De construto	É utilizado para mostrar que o instrumento de avaliação mede consistentemente conceitos teóricos.	Análise fatorial exploratória e confirmatória.
Convergente	Duas escalas ou dois métodos distintos convergem quanto ao significado do conceito operacionalizado no instrumento	Análise fatorial exploratória e confirmatória.
Discriminante	A amostra representativa de indicadores relacionados para com determinada dimensão latente não deve se relacionar com a natureza de outras construções.	Correlações entre dimensões latentes decorrentes das análises exploratórias e confirmatórias.
Preditiva	Outros indicadores são preditas em função das dimensões latentes validadas da escala.	Correlação momento produto de Pearson e análise de regressão.
Concorrente	Outros indicadores estão correlacionados com as dimensões latentes validadas da escala em um mesmo lapso temporal	Correlação momento produto de Pearson e análise de regressão.

Quadro 2: Tipos de validade requerida para com os instrumentos de avaliação da docência

4.3 Validação de construto

Validação é o processo usado para coletar evidências que apóiem os tipos de inferências a serem feitas a partir dos resultados de um teste. Segundo Moreira (1986), validação refere-se ao conjunto de procedimentos que dirá se o instrumento de avaliação tem ou não certo tipo de validade. O processo de validação é aceito quando se recorre aos

procedimentos analíticos de análise de dados para dotar o instrumento de certo grau de validade.

Nunnally (1978), descreve um processo de três passos usado para providenciar a validade de construto de um instrumento. São eles:

1. especificar o domínio de indicadores inerente ao construto;
2. a partir de investigação empírica e análises estatística, 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 altamente aceitas sobre o construto.

O primeiro passo sinalizado por Nunnally (1978) refere-se à relevância e representatividade do conjunto de indicadores de um instrumento, visando a tornar claro o conceito previamente definido. O estudo lógico do conjunto de teorias que facilitam a compreensão do construto e/ou opinião de especialistas na área ajudarão o pesquisador a estabelecer os indicadores que se correlacionam com o conceito investigado. Nessa etapa, procura-se estabelecer, através de procedimentos empíricos, o grau em que esse teste contém uma amostra representativa de indicadores do domínio de indicadores do construto. Para Ramos (1986), essa etapa é de importância capital. A definição do domínio do construto e de quais sejam seus indicadores definidos por hipótese, dependerá, em grande parte de todo o trabalho posterior de validação, assim como da importância das conclusões decorrentes do processo de validação.

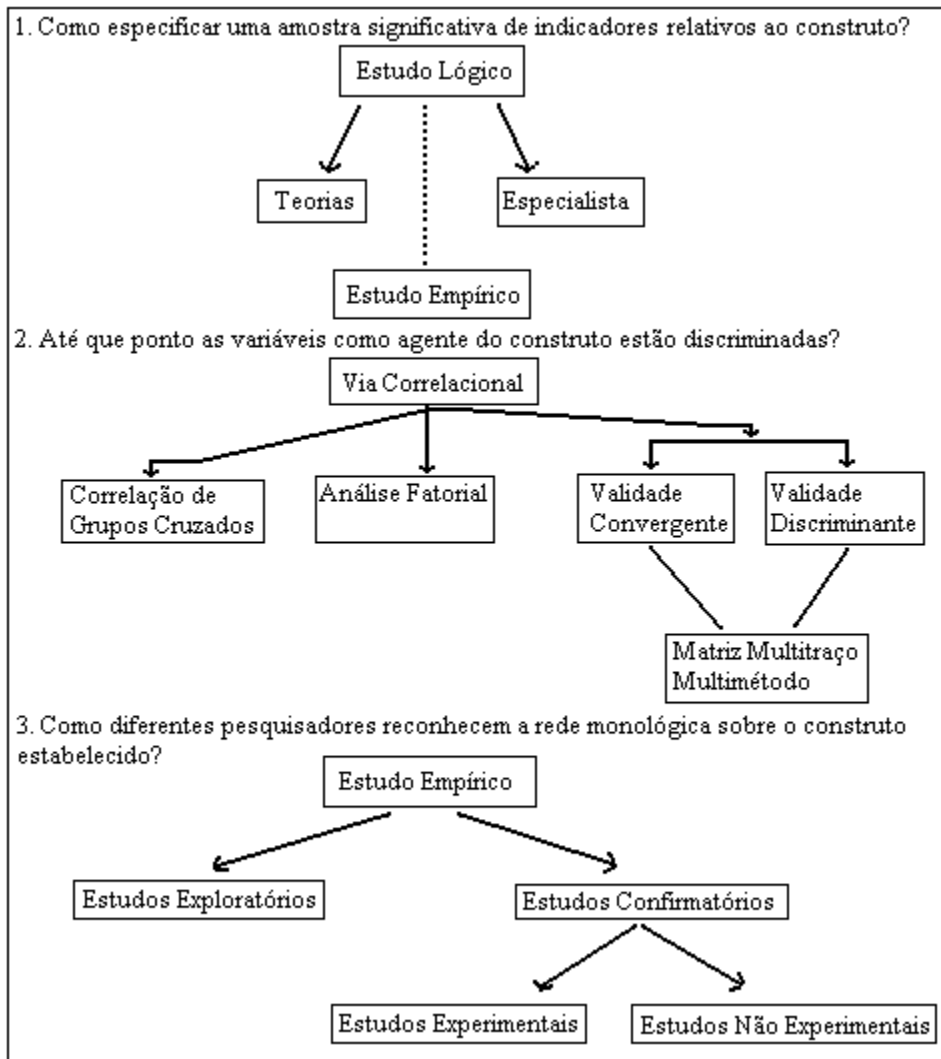
À medida que se avança no processo de validação de construto, é possível reformular a rede de relações e, dessa forma, melhorar a precisão dos indicadores do construto. Nessa etapa, utilizam-se, basicamente, métodos analíticos e correlacionais para mostrar que o conjunto de indicadores de determinado construto estão fortemente

intercorrelacionados entre si e não correlacionados com o domínio de indicadores que representam um outro construto, sendo que uma teoria de pano de fundo sustenta que os construtos não estejam correlacionados.

A ausência de correlações entre os construtos indica que os indicadores do construto estão bem discriminados; ou seja, entram com toda a força na definição do construto. Entre as diversidades de procedimentos estatísticos multivariados, que podem auxiliar o pesquisador na construção da teoria, está a análise fatorial, que, num contexto exploratório, auxilia a identificar os construtos que formam a base de um conjunto de indicadores distribuídos num dado instrumento de medida.

Na terceira etapa sinalizada por Nunnally (1978), deve ser testada a hipótese da legitimidade da representação dos construtos, que consiste em identificar a partir dos dados empíricos os construtos previamente operacionalizados no instrumento. Muitos são os procedimentos e técnicas que se pode utilizar nessa fase. Ramos (1986) apresenta que os estudos confirmatórios são os mais adequados para confirmar o construto e a sua medida. Entende-se por estudo confirmatório, não só aqueles estudos que se realizam em contextos experimentais ou causais, como também aqueles estudos cujos objetivos sejam o de validar as hipóteses teóricas originadas da rede monológica de um sistema de construtos (RAMOS, 1986). Para considerar um construto realmente válido, faz-se necessário acumular evidências empíricas, que sustentem a hipótese de que o conjunto de indicadores escolhidos para enfatizar um conceito de difícil acesso represente realmente, realmente, um construto hipotético (RAMOS, 1986).

O fluxograma 1, adaptado de Ramos (1986), apresenta os três passos sinalizados por Nunnally (1978), para assegurar a validade de construto de um instrumento e os principais procedimentos de análise de dados utilizados para acessar a sua validade.



Fluxograma 1: Aspectos importantes da validade de construto

4.4 Validação de construto – enfoque de Ramos

Para Ramos (1996), 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.

Ramos (1986), desdobra os três passos sinalizados por Nunnaly (1978) em cinco passos, que podem ser levados em consideração durante o processo de validação de um construto. São eles:

1. explicitação da teoria sobre o construto;
2. formulação das hipóteses derivadas da rede teórica e dos estudos exploratórios sobre o construto;
3. submeter as hipóteses à prova através de estudos confirmatórios;
4. avaliação dos resultados; e
5. formulação de novas hipóteses.

4.4.1 Explicitação da teoria sobre o construto

O processo de validação de um construto inicia-se com a formulação de definições minuciosas do construto, considerando os aspectos derivados da teoria psicológica, de pesquisas anteriores ou do processo de observação sistemática. Um construto é considerado um construto científico, quando o seu contexto teórico estiver bem explicitado. Usando a terminologia citada por Cronbach e Meehl (1967, *apud* RAMOS, 1987), um construto está bem explicitado, quando se define a rede monológica, ou seja, a quantidade mínima de associações ou proposições em torno do construto. Nessa etapa, os estudos exploratórios ajudarão o pesquisador na análise do domínio relevante para os indicadores do construto.

Estudos exploratórios são aqueles que auxiliam o pesquisador na ausência de um contexto teórico bem explicitado a definir as hipóteses específicas sobre a natureza de um construto teórico (RAMOS, 1986). Nessa fase, as técnicas multivariadas de análises, tais

como análise fatorial exploratória, análises de conglomerados, análises de correspondências e outras, são poderosas ferramentas de análises de dados que ajudarão na definição dos indicadores do construto.

4.4.2 Formulação das hipóteses derivadas da rede teórica e dos estudos exploratórios sobre o construto

Nessa etapa, o pesquisador procura estabelecer a rede teórica de relações do construto e distribuí-la adequadamente em um instrumento de medida. Cada construto, com seus múltiplos indicadores e seus interrelacionamentos, constitui em uma hipótese a ser testada por técnicas exploratórias de análise de dados. A submissão do instrumento a uma amostra randômica de indivíduos possibilitará dados que serão tratados por modelos exploratórios, objetivando rejeitar ou aceitar as hipóteses construídas.

Entre os modelos exploratórios, o mais amplamente utilizado durante o processo de validação de construto é o modelo de análise fatorial exploratória. Esse modelo estabelece que o pesquisador deve assumir que as respostas dos alunos aos indicadores do instrumento serão determinados por uma relação linear desses indicadores com os construtos, fato esse que nem sempre reflete a realidade.

Existe, ainda, um outro problema relacionado com estrutura fatorial escolhida, uma vez que a análise fatorial possibilita muitos fatores para uma mesma escala. Diante dessas dificuldades, não se tem a certeza da precisão do procedimento. Assim, Ramos (1986) recomenda que a hipótese a ser testada seja submetida a estudos não experimentais, preferencialmente do tipo confirmatório. Ainda, segundo Ramos (1987), das técnicas de estudo confirmatórios, a análise fatorial confirmatória ocupa papel de destaque durante o processo de validação de construto. A verificação da hipótese de legitimidade da representação dos construtos, pelo uso do modelo de análise fatorial confirmatória, procura

identificar, nos dados empíricos, um certo grau de validade dos construtos previamente operacionalizados no instrumento.

4.4.3 Submetendo as hipóteses à prova através de estudos confirmatórios

Segundo Ramos (1986), quando os métodos confirmatórios apóiam as hipóteses que se derivam da rede, o construto está se estabelecendo. Por outro lado, os dados podem discordar com as derivações da rede pelas seguintes razões:

1. os indicadores propostos como medidas do construto não o fazem adequadamente;
2. a rede teórica que originou as hipóteses é incorreta; e
3. o projeto experimental ou correlacional de validação fracassou, por não ter conseguido provar adequadamente a(s) hipótese(s).

4.4.4 Avaliação e integração dos resultados

Nessa etapa, a hipótese testada pode ser confirmada pelos índices decorrentes do processo de análise fatorial confirmatória, quais sejam: os índices de ajuste global do modelo, os índices de ajuste do modelo de medida e os índices de parcimônia. Quando esses indicadores da qualidade do ajuste se apresentarem conforme os padrões estabelecidos, as presenças de fontes sistemáticas de erro de medida não são considerados significativos. A consistência interna das dimensões latentes é um indicador útil que indica quão bem as medidas estão isentas de erros aleatórios.

A fase de avaliação de resultados, provenientes de distintos estudos realizados sobre as distintas hipóteses derivadas da rede total, exige um esforço lógico e teórico de integração, que deve realizar-se a partir da própria teoria que deu origem ao construto (RAMOS, 1996). Nessa etapa, exige-se uma análise de todos os resultados provenientes da

submissão do instrumento a contextos diversos. A convergência desses resultados agrega mais evidências a favor da validade do construto.

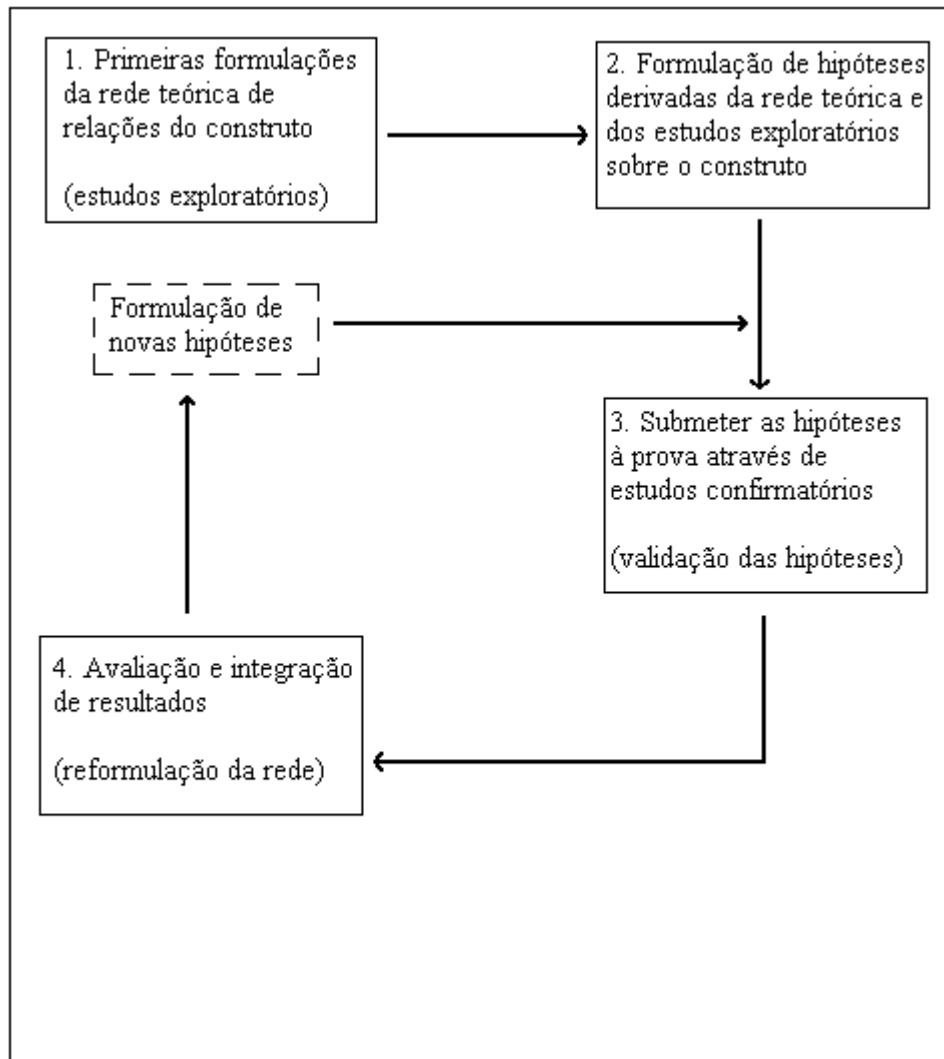
O resultado final dessa fase de avaliação e/ou integração de resultados é a consolidação das distintas hipóteses originadas da rede teórica ou a reformulação da própria rede teórica, no caso de inconsistência nos resultados das investigações.

4.4.5 Formulação de novas hipóteses

É uma volta ao segundo passo do processo de validação do construto e surge como resultado do processo de avaliação, que consolida o processo ou reformula a própria rede do construto. Qualquer alteração, na rede, vem a se constituir uma nova hipótese que deve ser investigada através de projetos de estudos confirmatórios. Caso o pesquisador não tenha clareza suficiente para constituir as novas hipóteses com relativa precisão, faz-se necessário a recorrência de estudos exploratórios. Os novos resultados alcançados devem ser avaliados e integrados entre si, com a conseguinte modificação ou consolidação do construto (RAMOS, 1986).

A aplicação dessa abordagem de validação, de acordo com Ramos (1987), é recomendada quando os dados de que se dispõe procedem da aplicação de um instrumento estruturado de medida e quando se pretende validar a estrutura fatorial inicial, bem como outras hipóteses estruturais alternativas.

O fluxograma 2, apresenta os passos sugeridos por Ramos (1986) para o processo de validação de construto.



Fluxograma 2: Passos do processo de validação de construto.

Pesquisas que tratam de validação de instrumentos de medida na área educacional, tais como os trabalhos de Marsh (1991), Ramos (1997), Capelleras e Veciana (2000), Greimel-Fuhrmann e Geyer (2003), Medina (2002), Mateo e Fernández (1992), Camisón, Gil e Roca (1999), utilizaram amplamente a metodologia de análise fatorial confirmatória, para comprovar as propriedades fundamentais de seus respectivos instrumentos de avaliação, quais sejam, as propriedades de confiabilidade e de validade.

4.5 Pesquisas sobre a validade da avaliação dos docentes pelos discentes

O processo de validação do instrumento de medida inicia-se com a análise lógica-teórica dos indicadores do instrumento, seguido de apoio empírico para as diversas dimensões ou conceitos teóricos, que esses indicadores unidimensionais estão medindo. A validade é muito mais difícil para se medir do que a confiabilidade. Segundo Cashin (1995), essa dificuldade é decorrente da falta de acordo sobre a definição de “ensino eficaz”, desse fato decorre a falta de um critério único para a sua avaliação de todo aceite. Devido a essa falta de critério, e por necessitar de uma avaliação da qualidade do ensino ministrada por seus professores, as universidades optam pela construção de instrumento de medida válido e preciso ou fidedigno, que possibilite avaliar através das opiniões dos alunos as diversas dimensões da qualidade do ensino, consideradas relevantes pela instituição.

Para que o instrumento usado não seja criticado pela baixa validade de conteúdo, é preciso acordo por parte da comunidade universitária sobre quais aspectos relevantes da qualidade do ensino esse instrumento deve medir. Quando a qualidade do ensino for medida em função do aprendizado do aluno, variáveis que dizem respeito ao conceito de professor efetivo e do quanto os alunos aprenderam com esse professor, devem fazer parte do instrumento. O instrumento pode ser projetado para refletir precisamente as opiniões dos alunos sobre a qualidade da instrução, indiferentemente se as avaliações reflitam o que os discentes aprendem (ABRAMI *et al.* 1990). Especificando o que o instrumento deve medir, a questão de validade fica evidenciada quando os indicadores unidimensionais do instrumento medem consistentemente as distintas dimensões da qualidade do ensino.

A validade dos níveis de satisfação dos alunos sobre a qualidade instrucional dos professores foi severamente questionada nos anos 70. A visão estreita em correlacionar somente as notas dos alunos com as avaliações dos professores fez com que a avaliação fosse

seriamente questionada. Para Snyder e Clair (1976), Worthington e Wong (1979), (*apud* GREENWALD, 1997), um professor pode obter uma avaliação boa, simplesmente atribuindo notas altas aos seus alunos, portanto fica comprometida a validade das avaliações dos alunos. No entanto, de acordo com Greenwald (1997), os experimentos, conduzidos pelos pesquisadores citados acima, nos quais os professores manipularam as notas dos alunos para cima e para baixo, para observar os efeitos nos níveis de satisfação dos alunos de maneira correspondente, foram revistos por diversos pesquisadores, concluindo que esses são coletivamente falhos. Isso significa que o aluno reconhece quando obtém uma nota acima de seu nível de aprendizado, logo o discente não troca o seu nível de satisfação para com determinado professor em função de melhores notas.

A partir de 1980, as pesquisas sobre a validade das avaliações dos alunos têm sido mais na forma de demonstrar que as avaliações dos alunos estão correlacionadas com outras variáveis importantes do ensino eficaz, tais como a interação do professor com a turma, a clareza expositiva do professor, conteúdo do professor, forma de avaliação e outras. Demonstrem, também, que estão menos correlacionadas com outras variáveis, tais como: o tamanho da classe, notas reais obtidas ou esperadas, motivação inicial do aluno etc. Dentro dessa perspectiva, a abordagem de validação de construto tem sido amplamente utilizada (GREENWALD, 1997).

Segundo Greenwald (1997), três tipos de estudos têm sido amplamente utilizados, para apoiar a validade de construto dos níveis de satisfação dos alunos com relação ao ensino recebido: estudos de Validade de Multiseção, proposto por Remmer, Martin e Elliot (1949), estudos de Modelos de Equação Estrutural (JÖRESKOG, 1969), estudos da Matriz Sintética de Multitraço-Multimétodo (CAMPBELL; FISKE, 1959).

A busca por validade das avaliações dos alunos através de estudos de validade de Multiseção vem produzindo evidências de que existem dimensões da qualidade do ensino que

são correlacionadas com o aprendizado dos alunos, evidenciados através das notas. O emprego dos modelos de equação estrutural e da matriz Multitraço-Multimétodo possibilita demonstrar que o instrumento mede algo concreto. Todos os três tipos de estudo podem ser utilizados para a demonstração da validade convergente das avaliações dos alunos.

4.5.1 Validade de construto – experimento de multiseção

É importante, também, para a validade dos níveis de satisfação dos alunos sobre o ensino recebido levar em consideração que os alunos dos melhores professores deveriam aprender mais e, conseqüentemente, obter notas altas nas provas e/ou trabalhos, enquanto os alunos dos piores professores deveriam aprender pouco e, conseqüentemente, obter notas baixas. Segundo d'Apollonia e Abrami (1997), um dos métodos para examinar esse critério é o estudo de validade de multiseção.

Apoio para o uso de experimentos de multiseção é dado por Marsh (1984). Esse autor coloca que é difícil correlacionar os níveis da satisfação dos alunos nas distintas dimensões do instrumento com as notas obtidas, uma vez que notas de provas de disciplinas diferentes, normalmente, não podem ser comparadas. O problema pode ser resolvido com os experimentos de multiseção, nos quais grupos diferentes de alunos, cursando a mesma disciplina, recebem os mesmos materiais e diferentes professores são designados para cada multiseção.

No estudo da validade de multiseção ideal, (a) existem muitas seções para um grande curso de multiseção; (b) os alunos são randomicamente designados para as seções ou no mínimo matriculam-se sem qualquer conhecimento sobre as seções ou quem irá examiná-los; (c) existem medidas de pré-teste disponíveis que correlacionam substancialmente com a performance dos alunos individuais no final da disciplina; (d) cada seção é ensinada complementarmente por um instrutor separado; (e) cada seção tem o mesmo esboço do curso, livros textos, objetivos do curso, e o exame final; (e) o exame final é construído para refletir os objetivos comuns por alguns especialistas que realmente não ensinaram em quaisquer umas das seções e, se houver um componente subjetivo, é avaliado por uma pessoa externa (MARSH, 1984, p. 720).

Segundo Marsh (1984), a evidência empírica para a validade dos níveis de satisfação dos alunos é alcançada quando as seções que demonstraram maiores níveis de satisfação com o ensino recebido, também, são as seções cujos alunos obtiveram, nos exames finais padronizados, as melhores notas.

Em virtude desse método de validação ter sido usado amplamente, com muitos formulários de avaliação do aluno sob diversas condições, ele providencia a evidência mais generalizável para a validade dos níveis de satisfação dos alunos (D'APOLLONIA; ABRAMI, 1997).

Cohen (1981), realizou uma meta-análise de todos os estudos de multiseção que não apresentaram falhas metodológicas. Usando as notas dos alunos no exame final de 68 cursos de multiseção e os vários itens de avaliação dos alunos, reportou as seguintes correlações apresentadas entre parênteses:

- 1) conquista ou aprendizado dos alunos (0,47);
- 2) avaliação total do curso (0,47);
- 3) avaliação global do professor (0,44);
- 4) dimensão: habilidade do professor (0,50);
- 5) dimensão: estrutura do professor (0,47);
- 6) dimensão: relacionamento do professor (0,31); e
- 7) dimensão: interação do professor (0,22).

Na meta-análise realizada por Feldman (1989) sobre os estudos de validade de multiseção, as notas finais dos alunos “estavam” razoavelmente correlacionadas com os níveis de satisfação dos alunos (correlações apresentadas entre parênteses), quanto às seguintes variáveis:

- 1) conquistas ou aprendizado (0,46);
- 2) preparação do curso/disciplina (0,57);

- 3) clareza dos objetivos (0,35);
- 4) compreensibilidade (0,56);
- 5) conhecimento da matéria (0,34);
- 6) disponibilidade do professor (0,36);
- 7) respeito pelos alunos (0,23); e
- 8) incentivo à discussão (0,36).

As correlações, reportadas por Cohen (1981) e Feldman (1989), apóiam a conclusão de que as classes, nas quais os alunos atribuíram ao professor avaliações mais elevadas, tendem a ser aquelas em que os alunos alcançaram maiores conquistas. Assim, as avaliações dos alunos são medidas válidas da eficácia do ensino. Por outro lado, as correlações encontradas não são altas, uma vez que o melhor aprendizado do aluno depende em grande parte de suas habilidades e não das características do professor (CASHIN, 1995).

As sete dimensões do instrumento de avaliação de Frey *et al.* (1975) foram validadas através de experimentos de multiseção por alunos de três universidades: Purdue com nove seções, North Dakota State com cinco seções e Northwestern com 12 seções. As sete dimensões do instrumento *ENDEAVOR*: clareza da apresentação, carga de trabalho, atenção pessoal, discussão em classe, organização/planejamento, forma de avaliar, conquistas dos alunos, respectivamente apresentaram as seguintes correlações médias: 0,58; -0,28; 0,38; 0,36; 0,51; 0,30; 0,59 com as notas dos alunos no exame final. Como seis das correlações obtidas superam a 0,29, há razoáveis evidências empíricas para a validade do instrumento *ENDEAVOR*, ou seja, o instrumento apresenta seis dimensões correlacionadas com a efetividade do ensino (aprendizado do aluno).

Centra (1977), investigou a validade das avaliações dos alunos em 72 projetos de multiseção, envolvendo sete disciplinas diferentes. Para duas disciplinas, química e biologia, respectivamente com sete e sete seções, os alunos foram designados randomicamente,

enquanto que para as outras cinco disciplinas física, biologia, matemática, psicologia e química, respectivamente, com 22, 13, 8, 7 e 8 seções, os alunos foram designados de forma não aleatória. Nesse estudo, Centra correlacionou as notas dos alunos, no exame final, com os indicadores da qualidade de ensino do docente, apresentadas abaixo, obtendo as seguintes correlações medianas expostas entre parênteses:

- a) efetividade do ensino (0,60);
- b) valor da disciplina para o estudante (0,61);
- c) relacionamento professor/aluno (0,30);
- d) objetivos e Organização da disciplina (0,49);
- e) indicação de leitura (0,35);
- f) dificuldade da disciplina e carga de trabalho (-0,13);
- g) exames (0,31);
- h) qualidade das leituras (0,47); e
- i) esforço dos estudantes (0,09).

Os resultados de Centra (1977), relatam que as avaliações dos alunos sobre as variáveis “demanda de esforços pelos estudantes” e “dificuldade da disciplina/carga de trabalho” não estão correlacionadas com a *performance* dos alunos nos exames finais.

D’Apollonia e Abrami (1996 *apud* D’APOLLONIA; ABRAMI, 1997) relatam que a sua meta-análise extraiu todos os coeficientes de validade de 43 estudos de validade de multiseção. Eles reportaram que o coeficiente de validade médio dos níveis de satisfação dos alunos sobre a instrução recebida geral nesses estudos foi de 0,33, estendendo-se a partir de 0,29 a 0,37, com confiabilidade de 95%.

Para Cashin (1988, 1955), ao se analisar correlações de validade, percebeu que correlações acima de 0,7 não são freqüentes em medidas educacionais, especialmente se for estudado fenômenos complexos, tais como o aprendizado do aluno. Esse autor sugere que os

coeficientes de validade para os níveis de satisfação dos alunos entre 0,0 e 0,29, até quando estatisticamente significativos como inexistentes, coeficientes de validade entre 0,30 e 0,49 são classificados como sendo razoáveis e que as correlações entre 0,5 e 0,7 são tidas como boas.

Usando a escala de classificação acima e os resultados reportados por Frey *et al.* (1975), Centra (1977), Cohen (1981) e Feldman (1989), D'Apollonia e Abrami (1996), é possível concluir que os níveis de satisfação dos alunos, sobre as diversas dimensões do ensino eficaz, é válido para acessar o construto “efetividade do ensino”, ou seja, as seções, nas quais os alunos manifestaram maiores níveis de satisfação, nas diversas dimensões do instrumento, correspondiam às seções onde os discentes apresentaram maiores *performances* no exame final.

Marsh (1984), apresenta os seguintes argumentos para os resultados positivos, nos estudos de validade de multiseção, quando as notas do curso, conhecidas ou esperadas, e a *performance* no exame final, estão substancialmente correlacionadas com as avaliações dos alunos:

- 1 um ensino mais efetivo, que proporcione maior aprendizado, logo o professor é melhor avaliado;
2. satisfação crescente do aluno com notas maiores as quais fazem atribuir níveis de satisfação mais elevado ao professor, independente dos níveis de aprendizado do aluno; e
3. diferenças individuais nas características dos alunos, como, por exemplo, interesse anterior pela matéria, motivação e habilidade, que afetam ambos, notas obtidas pelos alunos e avaliação dos professores feita pelos alunos.

A primeira hipótese fornece o suporte empírico para a validade das avaliações dos alunos como medida da efetividade do ensino, a segunda pode ser interpretada como um viés

indesejável nas avaliações, enquanto a terceira é o efeito de variáveis prognósticas que são precisamente refletidas pelas avaliações dos alunos (MARSH, 1984).

1.5.1 Validade de construto – modelos de equação estrutural

Os modelos de equação estrutural com variáveis latentes e erros de medida apresentado por Jöreskog (1978), utilizadas como estratégia de modelização confirmatória por Marsh (1991), Mateo e Fernández (1992), Ramos (1997), Val (1994), Greimel-Fuhrmann (2003), Capelleras e Veciana (2002) e Medina (2002), têm como objetivo fornecer o suporte empírico para a validade de construto do instrumento de avaliação da docência, ou seja, procura avaliar a validade de uma série de indicadores como medida de outras variáveis de interesse teórico (variáveis latentes ou construtos).

Para Ramos (1987), o modelo de análise fatorial confirmatória é muito útil em processos de validação de construto, especialmente quando os dados que se dispõe são resultados da aplicação de um instrumento estruturado de medida e que se pretende validar determinada hipótese estrutural decorrente da análise lógica do conteúdo ou validar outras hipóteses estruturais alternativas.

Ramos (1997), utilizou a técnica estatística multivariada de análise fatorial confirmatória na validação do construto competência docente do professor universitário. O estudo teve como objetivo encontrar um modelo de medida da competência docente do professor universitário, sobre a base de uma escala de valoração das características da docência desse professor. Os resultados das análises fatoriais exploratórias efetuadas, *a priori*, possibilitaram a construção de seis modelos que passaram a constituir seis hipóteses alternativas. Utilizando o *LISREL* (Relações Estrutural Linear) como modelo de análise e dos indicadores da qualidade de ajuste, concluiu-se que um modelo de medida definido por seis

dimensões correlacionadas (Programação/Organização do Ensino; Domínio do Conteúdo/Clareza Expositiva; Motivação para o Aprendizado/Crescimento do Interesse do Aluno; Interação com a Classe; Atenção Individual ao Aluno e Avaliação/Exames) são atributos do professor indispensáveis para uma docência de qualidade. Todas as dimensões citadas são constituídas por cinco indicadores que caracterizam a atividade do docente.

Mateo e Fernández (1992), desenvolveram um questionário destinado a medir a competência evidenciada pelos professores e suas capacidades para motivar os discentes. O instrumento foi elaborado e validado pela metodologia de análise fatorial confirmatória, originalmente na Espanha. Tal estudo consta de 31 indicadores que saturam unidimensionalmente em oito construtos teóricos ou dimensões. As oito dimensões foram as seguintes:

dimensão I: estilo docente (cinco indicadores);

dimensão II: estruturação dos Conteúdos (três indicadores);

dimensão III: clareza Expositiva (três indicadores);

dimensão IV: competência Acadêmica (dois indicadores);

dimensão V: habilidades Motivacionais e de Interação (dois indicadores);

dimensão VI: requisitos docentes (três indicadores);

dimensão VII: equilíbrio Avaliativo (três indicadores); e

dimensão VIII: critérios Gerais (dois indicadores).

Marsh (1991), submeteu o instrumento *SEEQ* (Estudantes Avaliam a Qualidade Educacional) à abordagem de validação de construto, com o propósito de questionar a unidimensionalidade do *SEEQ*. Quatro modelos foram testados, utilizando duas amostras randômicas de médias de 500 classes, que originaram duas sub-amostras. A sub-amostra um foi destinada à calibração, enquanto a sub-amostra dois à validação cruzada. A análise fatorial confirmatória (AFC) das respostas do *SEEQ* forneceu o suporte empírico para a

multidimensionalidade do *SEEQ*, nove fatores de primeira ordem. Os nove fatores de primeira ordem do *SEEQ* foram submetidos à análise fatorial confirmatória hierárquica. A análise cuidadosa dos modelos estimados possibilitou concluir que a maioria da variância residual padronizada do escore verdadeiro dos fatores de primeira ordem não foi adequadamente explicada pelos fatores de segunda ordem. O estudo demonstra que as respostas do *SEEQ* não podem ser adequadamente explicadas por um ou mais fatores de segunda ordem.

Barton e do Andrew (1994), examinaram a validade fatorial do instrumento TES (Avaliação da Efetividade do Ensino), elaborado para medir a efetividade do ensino em programas de graduação de professores. A amostra formada por 390 classes de 33 universidades americanas foi randomicamente dividida em duas sub-amostras. A primeira sub-amostra, composta pelas informações contidas em 195 unidades de análise, foi submetida à análise fatorial exploratória, enquanto que a segunda sub-amostra, formada pelas informações contidas em 195 avaliações, foram submetidas à análise fatorial confirmatória. As análises exploratórias e confirmatórias convergem a três dimensões consistentes e intercorrelacionadas do TES. Segundo os autores, a quantia alta das correlações das três sub-escalas do TES pode ser um indicativo da existência de um fator único, global, de ordem elevada da efetividade do ensino, ou decorre do fato de que se um docente for bem avaliado em um determinado fator, será avaliado da mesma forma nos demais.

Capelleras e Veciana (2000), utilizaram a abordagem da análise fatorial confirmatória, para obter o suporte empírico da validade de construto de um instrumento destinado a medir a qualidade do serviço no ensino universitário. Dois modelos foram testados. O primeiro com todas as variáveis associadas a um único fator global e o segundo, com cinco fatores identificados pela análise lógica do conteúdo e pela técnica de análise fatorial exploratória. Todos os indicadores de ajuste absoluto, incremental e de parcimônia,

indicam que os dados da amostra de 811 questionários se ajustam bem ao modelo proposto de cinco fatores latentes. O estudo mostrou que os dados sustentam a hipótese da validade de construto-convergente. Todas dimensões da escala estão significativamente correlacionadas com as variáveis critério “qualidade global percebida pelos discentes” e “satisfação dos alunos na universidade”. A tabela 8, apresenta as dimensões do instrumento e os coeficientes de correlação dessas dimensões com a qualidade global percebida pelos alunos e satisfação desses com a universidade.

Tabela 8: Coeficientes de correlação entre as dimensões do instrumento e qualidade global e satisfação.

Dimensão	Correlação com a qualidade global	Correlação com a satisfação dos alunos
Atitudes e comportamento dos professores	0,580	0,470
Competência dos professores	0,399	0,400
Projeto pedagógico do curso	0,419	0,319
Instalações e equipamentos	0,395	0,348
Organização do ensino	0,301	0,139

Fonte: Capelleras e Veciana, 2000.

O resultado reportado na tabela 8 mostra que todas as correlações são positivas e diferentes de zero (p -valor $< 0,01$). As correlações encontradas fornecem o suporte empírico de validade concorrente, ou seja, a escala apresenta capacidade para predizer outras variáveis.

Os primeiros passos para efetuar a avaliação do desempenho dos docentes a partir da opinião dos alunos por instrumento tecnicamente correto, também, foi dado por Medina (2002). Duas hipóteses derivadas da análise lógica dos conteúdos e das diversas análises

fatoriais exploratórias foram propostas e testadas através da técnica de análise fatorial confirmatória.

A primeira hipótese, submetida à prova, afirma que, segundo os alunos, o bom desempenho do docente depende das suas habilidades para direcionar a sua disciplina, conforme os objetivos do curso (ênfase do curso), manter os alunos motivados (estilo interpessoal), domínio da matéria, respeito aos alunos, se utiliza critérios justos para avaliar e qualificar, se prepara a sua aula e cumpre o programa e dinâmico frente à classe. O modelo supõe que essas sete dimensões sejam correlacionadas.

A segunda hipótese a ser testada afirma que o bom desempenho dos docentes depende dos níveis de satisfação dos alunos em duas dimensões de primeira ordem, quais sejam: avaliação / qualificação e ênfase do curso e em duas dimensões de segunda ordem, denominadas: competência para o ensino que engloba os fatores de primeira ordem “organização/preparação da classe” e “domínio da matéria” e estilo de interação com os alunos que engloba os fatores de primeira ordem “respeito com os alunos” e “motivação e estilo pessoal”. Os indicadores de bondade de ajuste foram estimados para os dois modelos, porém os indicadores do modelo II são mais significativos em termos estatísticos. Quanto à consistência interna das dimensões do modelo II, todos excederam a 0,7 exceto a dimensão “avaliação e qualificação”, que o autor recomenda revisar a conformidade da escala, incluindo um número maior de indicadores.

Greimel-Fuhrmann e Geyer (2003), avaliam os fatores que determinam a avaliação dos estudantes do ensino médio sobre os seus professores de contabilidade. O instrumento é composto por nove itens de avaliação global e por 42 itens que cobrem os diferentes aspectos do bom ensino. A análise empírica da multidimensionalidade dos diversos itens que cobrem os diferentes aspectos do comportamento dos docentes através de modelo de equação estrutural resulta em três dimensões correlacionadas. Essas três dimensões relatam

que a avaliação global dos estudantes sobre os seus professores de Contabilidade depende da sua capacidade de orientar a disciplina, da sua capacidade para orientar os estudantes no curso e da capacidade do professor para o gerenciamento da classe, com coeficientes de correlação bivariada, respectivamente iguais a 0,707; 0,418; 0,263, todos estatisticamente significante ao nível de 1%. O Modelo de equação estrutural testado apresenta índices incrementais de ajuste superiores a 0,9, adequado sob o ponto de vista estatístico. Os resultados fornecem o suporte empírico de validade de construto-convergente.

Cajide Val (1994), submeteu a matriz de correlação dos nove fatores de primeira ordem, correspondentes às características da qualidade docente à análise fatorial confirmatória de segunda ordem, com a finalidade de obter uma estrutura fatorial simplificada das dimensões da qualidade do docente universitário. Dois métodos de estimação foram empregados: o método dos mínimos quadrados de dois estágios e o método de máxima verossimilhança. Os métodos de resolução convergiram para uma solução fatorial de três fatores de segunda ordem altamente correlacionados, com indicador incremental de ajuste menor que 0,90 (AGFI) e indicadores absolutos de ajuste, equivalente à cota mínima de aceitação (0,907). O primeiro fator, denominado interação, é representado pelos seguintes fatores de primeira ordem: interação com os alunos, avaliação, comunicação, metodologia e recursos e gestão/qualidade. O segundo fator está centrado no professor e representa a sua habilidade para a investigação com qualidade e para com o desenvolvimento do programa. O terceiro fator está relacionado com a habilidade do professor para organizar e dominar a matéria. Como o modelo estimado está baseado em informação, *a priori*, sobre a natureza da estrutura dos dados em forma de uma teoria específica ou hipóteses e apresenta indicadores da qualidade do ajuste razoáveis, é possível argumentar que os dados fornecem o suporte empírico para a validade de construto do instrumento.

Os modelos de análises fatoriais confirmatórias é, na atualidade, o método mais potente para a investigação da validade de construto das escalas de medida. A sua empregabilidade decorre do fato de que é possível impor umas séries de restrições de partida, suficientemente, apoiada na teoria, tais como: as variáveis observadas estão afetadas por outros fatores, que os indicadores apresentam erro de medida, que os erros de medida estão correlacionados, que as dimensões latentes estão correlacionadas, que os indicadores possuem carga sobre um único fator latente e, a seguir, efetuar a modelagem. Além dessas vantagens, existem uma série muito grande de indicadores de ajustes absolutos, incrementais e de parcimônia que podem ser usados conjuntamente para a investigação da qualidade do ajuste estabelecido.

1.5.2 Validade de construto – matriz de multitraço-multimétodo

A matriz de multitraço-multimétodo desenvolvida por Campbell e Fiske (1959) é uma das técnicas mais usuais no processo de validação de construto. Esse método integra simultaneamente o estudo de validez de construto convergente e discriminante. A validade convergente é evidenciada pelas altas correlações entre as dimensões de distintos instrumentos independentes que se supõe medirem o mesmo conceito. A validade divergente ou discriminante é alcançada pela inexistência de correlação significativa entre dimensões com as quais teoricamente deveria diferir. Quando a validade discriminante é evidenciada, cada dimensão está bem concebida enquanto distinta (TOURÓN, 1989).

Touron (1989), submeteu as nove dimensões do instrumento *SEEQ* e as sete dimensões do instrumento *ENDEAVOR*, questionários destinados à avaliação da eficácia do docente, à matriz de multitraço-multimétodo. As dezesseis dimensões desses dois instrumentos constituem os multitraços, enquanto os dois instrumentos são os multimétodos.

Os coeficientes de validade convergente, apresentados na tabela 9, correspondem às dimensões de ambos instrumentos que se supõem medirem o mesmo conceito. Como se pode constatar, todas as correlações são suficientemente altas e significativamente diferentes de zero. Todos os pressupostos que favorecem o apoio empírico para a validade de construto-convergente e discriminante foram amplamente estabelecidos. Estes resultados evidenciam a alta qualidade técnica desses instrumentos.

Tabela 9: Coeficientes de convergência da matriz de mutitraço-multimétodo dos instrumentos de avaliação de Marsh (1982) e de Frey *et al* (1975).

Dimensões dos instrumentos SEEQ *ENDEAVOR	Correlações (Coeficientes de Convergência)
Interação com o grupo * Discussão em classe	0,93
Aprendizagem/valor * Aprendizado dos alunos	0,86
Carga de trabalho/dificuldade * Carga de trabalho	0,82
Exames/qualificação * Qualificação	0,80
Relacionamento individual * Atenção pessoal	0,81
Organização/claridade * Claridade na exposição	0,78
Organização/claridade * Organização/planejamento	0,71

Fonte: Tourón (1989).

Marsh e Hocevar (1989), apresentam o suporte empírico para a validade de construto do instrumento *SEEQ*. Esses autores analisam as avaliações dos professores universitários que lecionaram a mesma disciplina, no mínimo quatro vezes em um período de

quatro anos. As quatro análises fatoriais para as respostas médias das 316 classes da mesma disciplina, foram feitas separadamente, respectivamente, ao período em que foi oferecida. Tais análises convergiram quanto à estrutura fatorial do instrumento, ou seja, todas as análises fatoriais exploratórias replicaram as nove dimensões do *SEEQ*.

Os autores submeteram as matrizes de correlações entre os fatores desses quatro estudos à duas análises mais criteriosas: matriz de multimétodo-multitraço e a análise fatorial confirmatória. Os estudos da matriz de multitraço-multimétodo, em que as nove dimensões do *SEEQ* são os múltiplos traços e as quatro ofertas da mesma disciplina são os múltiplos métodos, revelou um coeficiente de validade convergente médio igual a 0,68. Os resultados evidenciam a validade de construto do *SEEQ*.

A desvantagem da aplicabilidade da matriz de multitraço-multimétodo para evidenciar a validade de construto-convergente e discriminante, está no fato de requerer dois instrumentos distintos, construídos de forma independentes, para medirem os mesmos conceitos e, posteriormente, valorá-los por duas amostras independentes de alunos. Sendo o ensino eficaz um construto multidimensional, ou seja, pode-se obtê-lo de diversas formas, as universidades optam por construir instrumentos de forma independente, o que torna a metodologia pouco usual para validar os construtos operacionalizados nos instrumentos de avaliação da docência. Por outro lado, pode ser útil, caso o mesmo instrumento seja valorado por uma amostra de alunos e uma amostra de professores. De sua análise, pode-se concluir quando, alunos e professores reconhecem as mesmas dimensões da qualidade do ensino.

4.6 Outros critérios relacionados com a avaliação dos estudantes

A avaliação realizada pelos alunos universitários sobre a efetividade de seus professores tem sido correlacionada com as auto-avaliações dos professores, com as

avaliações dos colegas, com as avaliações de observadores externos e com a produtividade de pesquisa do professor com o propósito de obter evidências empíricas para a validade dos níveis de satisfação dos alunos (MARSH, 1984).

As auto-avaliações dos docentes, utilizando-se do mesmo formulário de avaliação do aluno ou de um outro, é de grande ajuda para o melhoramento da qualidade do ensino do professor. Por um lado, propicia ao docente um momento de reflexão da qualidade do ensino por ele ministrado, por outro, possibilita ao professor comparar a sua auto-avaliação com a avaliação dos alunos ou de outras fontes. O confronto das informações decorrentes de fontes diversas facilitará ao professor o estabelecimento de ações corretivas com vistas ao melhoramento do processo de ensino e de aprendizagem. Segundo Sanchez (1985), muitos pesquisadores utilizam o mesmo formulário de avaliação do aluno para a auto-avaliação do professor, de tal forma que se obtém uma validade convergente.

Nesse sentido, Marsh (1982c *apud* MARSH, 1984) apresenta um estudo em que os docentes de 81 disciplinas foram solicitados para avaliar seu próprio ensino, no mesmo instrumento de avaliação completado pelos alunos. As respostas médias dos alunos e professores foram submetidas a análises fatoriais exploratórias, as quais detectaram a mesma estrutura fatorial do SEEQ. O suporte empírico, para a validade de construto-convergente para as dimensões evidenciadas pelos alunos e pelos professores, foi obtido mediante o emprego da matriz de multitraço-multimétodo, que revelou os coeficientes de validade convergente, apresentada na tabela 10.

Tabela 10: Coeficiente de validade convergente das dimensões do instrumento *SEEQ*, evidenciados a partir das respostas dos alunos e dos professores.

Dimensões	Coeficiente de Validade Convergente
Aprendizado/Valor	0,46
Entusiasmo	0,54
Organização	0,30
Interação em Grupo	0,52
Relacionamento Individual	0,28
Profundidade dos Temas	0,42
Exames	0,17
Tarefas/Leitura	0,45
Trabalho/Dificuldade	0,69

Fonte: Marsh (1984).

Os coeficientes de validade convergentes reportados na tabela **10** mostram concordância significativa entre as avaliações dos alunos e as auto-avaliações dos professores. Esses resultados demonstram que professores e alunos reconhecem as mesmas dimensões da qualidade do ensino do instrumento de avaliação de Marsh e providencia um certo grau de validade convergente para os níveis de satisfação dos alunos.

Numa revisão da literatura, Feldman (1989) cita 19 estudos que correlacionam as auto-avaliações dos professores com as avaliações dos alunos. A correlação média entre as avaliações dos alunos e as auto-avaliações dos professores, nesses estudos, foi de 0,29. Marsh (1984) cita dez estudos que correlacionaram as auto-avaliações dos professores com as avaliações dos alunos. As correlações variaram entre 0,20 e 0,69. Feldman (1988) revisou 31 estudos que tratam da percepção que alunos e professores possuem sobre uma concepção apropriada do que é o ensino eficaz e reportou uma correlação de 0,71 entre a visão dos alunos e professores. Os níveis de correlações reportados por Feldman (1988, 1989) e Marsh

(1984) providenciam apoio adicional para a validade convergente dos níveis de satisfação dos alunos com relação ao desempenho de seus professores.

Os resultados acima evidenciam que alunos e professores reconhecem os atributos necessários para uma boa docência. Assim, esses atributos podem ser utilizados junto aos professores como argumentos favoráveis às avaliações dos alunos perante aqueles professores que ainda acreditam que os alunos não são capazes de realizarem uma avaliação justa da qualidade do ensino recebido.

Na avaliação dos docentes universitários pelos colegas, utilizada principalmente em universidades, todo trabalho do professor universitário, seja no âmbito da sala de aula, seja na pesquisa, seja na prestação de serviços para a comunidade, é avaliado por um outro professor, coordenador ou chefe de departamento. Em geral, quando os professores são avaliados pelos seus colegas, encontra-se correlação entre as avaliações feitas pelos colegas e as avaliações feitas pelos alunos, já que esses são os que providenciam informações para os colegas. As avaliações dos professores pelos colegas são, em geral, menos confiáveis, menos válidas do que as realizadas pelos alunos, são percebidas como mais ameaçadoras. Criam um ambiente ruim e são afetadas por outras variáveis não relacionadas com a eficácia do docente, tal como produtividade na pesquisa (SANCHEZ, 1995).

Morsh, Burgess e Smith (1956, *apud* MARSH, 1984) correlacionaram as avaliações dos alunos, rendimentos dos discentes, avaliações dos colegas e avaliações do supervisor num grande curso de multi-seção: as avaliações dos docentes pelos alunos correlacionaram-se com as suas conquistas, apoiando sua validade. As avaliações dos colegas e supervisores, embora significativamente correlacionadas, não estavam correlacionadas seja com as avaliações dos alunos ou com as suas conquistas. O estudo sugere que as avaliações dos colegas possam não ter valor como um indicador do ensino eficaz. Neste sentido, Centra (1979), French-Lazovich (1981), (*apud* MARSH, 1984) têm também falhado em identificar

estudos que fornecem apoio empírico para as avaliações dos colegas como um indicador do ensino universitário eficaz ou como um critério para as avaliações dos alunos. Por outro lado, Kulik e McKeachie (1975, *apud* CASHIN, 1988) relatam correlação entre avaliação dos professores feitas pelos alunos e pelos colegas, variando entre 0,48 e 0,69.

Webb e Nolan (1955, *apud* MARSH, 1984) relatam uma boa harmonia entre as avaliações dos alunos e as auto-avaliações dos professores, mas nenhum desses indicadores estava positivamente correlacionado com as avaliações do supervisor, que os autores indicaram para serem parecidos com as avaliações dos colegas. Em síntese, Centra (1979), Koon e Murray (1996), Marsh (1987), Murray (1980), (*apud* MARSH; ROCHE, 1997) citam que as avaliações dos colegas e administradores, baseados nas visitas em sala de aula, não são confiáveis, ou seja, as avaliações por colegas diferentes nem ao menos estão de acordo umas com as outras e não estão sistematicamente correlacionadas com as avaliações dos discentes sobre o ensino recebido ou outros indicadores do ensino eficaz.

Howard, Conway e Maxwell (1985) investigaram a validade de construto referentes aos seguintes métodos de avaliação da efetividade do ensino: avaliações de alunos, avaliação de colegas, avaliações feitas por alunos treinados, avaliações de ex-alunos e auto-avaliação dos professores. Os autores concluíram que as avaliações dos professores feitas por alunos e ex-alunos evidenciam coeficientes de validade substancialmente maiores da eficácia do ensino do que o fazem as avaliações de colegas, avaliações feitas por alunos treinados e auto-avaliação dos professores. McKeachie (1997), também cita que as avaliações dos alunos são a única fonte válida de dados sobre a eficácia do ensino. Nesse sentido, Marsh e Roche (1997) enfatizam que existe pouca evidência da validade de quaisquer outras fontes de dados.

4.7. Fontes possíveis de viés nas avaliações dos alunos

Especificamente, dois tipos de pesquisas têm pretendido invalidar a validade das avaliações dos alunos. A primeira é aquela que procura prova de efeitos de contaminação ou viés nas avaliações dos alunos. A segunda é aquela que procura por falhas metodológicas no processo de demonstrar correlações substanciais positivas entre as avaliações dos alunos e as diversas variáveis que indicam a eficácia do ensino.

Denomina-se “viés”, de acordo com Marsh (1984), as variáveis correlacionadas com as avaliações dos alunos que não estão correlacionadas com a eficácia do ensino. Por essa definição, as correlações entre as avaliações dos alunos e o tamanho da classe, ou se o interesse antecipado do aluno na disciplina não são “vieses”, porque é provável que os alunos em pequenas classes, ou classes de alunos que estejam interessados na matéria de fato aprendem mais, enquanto a variável brandura nas notas, ou seja, docentes atribuindo notas mais altas do que os alunos merecem, pode enviesar as avaliações dos alunos. Recomenda-se que se tenha o conhecimento prévio dessas variáveis, supostas contaminadoras das avaliações, dos alunos para que sejam investigadas, e dessa forma, controlá-las no processo de avaliação dos alunos e disseminação dos resultados. Segundo Marsh (1984), a validade de construto das avaliações dos alunos requer que as avaliações estejam relacionadas com as variáveis que são indicativas do ensino eficaz, mas não correlacionadas com variáveis supostas a enviesarem as avaliações dos alunos.

Marsh e Overall (1979, *apud* MARSH, 1984) questionaram aos membros docentes sobre quais variáveis de uma lista de 17, supõe-se enviesarem as avaliações dos alunos. Os docentes (percentuais entre parênteses) ressaltaram que a avaliação do aluno é influenciada pela:

a) dificuldade da disciplina (72%);

- b) severidade ou brandura nas notas (68%);
- c) popularidade do professor (63%);
- d) interesse prévio do aluno pela disciplina (62%);
- e) carga de trabalho na disciplina (60%);
- f) tamanho da classe (60%);
- g) razão para cursar a disciplina (obrigatória ou optativa) (55%); e
- h) nota média do aluno (53%).

Segundo Sanchez (1985), diversos estudos evidenciam que as variáveis “interesse prévio do aluno pela disciplina”, “carga de trabalho na disciplina”, “tamanho da classe”, “razão para cursar a disciplina” e “nota média do aluno” mostram que, conjuntamente, explicam entre 5% e 20% da variância total das avaliações dos alunos; enquanto que as variáveis “dificuldades da disciplina”, “severidade ou brandura nas notas”, “popularidade do professor”, “carga de trabalho na disciplina” explicam, conjuntamente, mas diretamente, em torno de 5% a 20% da variância total das avaliações dos alunos. O “interesse antecipado pela disciplina”, “notas esperadas” e talvez, “carga de trabalho/ dificuldade” parecem ser as variáveis de pano de fundo mais fortemente correlacionadas com as avaliações dos alunos (MARSH, 1984).

Segundo Marsh (1984), uma correlação positiva e significativa entre as variáveis enviesantes das avaliações dos docentes pelos discentes não é uma evidência definitiva de invalidez das avaliações. O apoio para uma hipótese de viés deve ser baseado numa abordagem de validação de construto. Tal abordagem requer que as variáveis “supostas vieses” das avaliações dos alunos sejam examinadas em estudos que estejam relativamente livres de falhas metodológicas e que sejam interpretadas em relação a uma definição específica de viés. Muitos estudos que tratam da validade das avaliações dos alunos estão mal interpretados, segundo Marsh (1984), pelas seguintes razões:

1. uma correlação não implica causalidade;
2. confunde-se significância estatística com significância prática;
3. não se leva em consideração a natureza multivariada das avaliações;
4. utilizam-se dados individuais e não as médias das classes;
5. não se replicam os resultados dos estudos em outros contextos;
6. falta uma definição clara do que seja uma fonte de invalidez; e
7. questões de conveniências de manipulações experimentais.

Marsh (1984), Cashin (1988, 1995), Marsh e Roche (1997) e Aleamoni (1999) apresentam uma síntese das principais pesquisas que tratam do relacionamento das variáveis que supostamente contaminam as avaliações dos docentes pelos discentes. Esses autores apresentam as variáveis:

1. **motivação do aluno** - Classes com um percentual de alunos, que tinham um interesse anterior pela matéria, tendem a avaliar o professor mais favoravelmente, apesar de que nem sempre esteja claro se o interesse existiu antes do início da disciplina ou se o interesse foi despertado pelo professor (MARSH; ROCHE, 1997).
2. **notas esperadas** - As notas médias das classes estão correlacionadas positivamente com as avaliações médias da classe de alunos sobre o ensino do professor. Isso não significa que as avaliações podem ser aumentadas pela estratégia de atribuir maiores notas para os alunos.
3. **razão para fazer o curso** - Disciplinas eletivas e aquelas com um percentual de alunos fazendo o curso por interesse geral tendem a avaliar melhor os professores (MARSH; ROCHE, 1997).
4. **carga de trabalho/dificuldade** - Segundo Centra (1993), Marsh e Dunkin (1992), (*apud* CASHIN, 1995) e Marsh e Overall (1979b, *apud* MARSH, 1984), existe relacionamento positivo entre carga de trabalho/níveis de dificuldade do curso e as avaliações dos alunos,

- ou seja, alunos atribuem melhores avaliações aos professores que exigem que aqueles estudem bastante. Carga de trabalho/dificuldade não parece constituir um viés para as avaliações dos docentes pelos discentes (MARSH, 1984).
5. **tamanho da classe** - Segundo Cashin (1995), embora haja uma tendência para classes menores avaliarem os seus professores mais favoravelmente, essa é uma correlação muito fraca. Para Marsh (1984), o tamanho da classe se correlaciona moderadamente com algumas dimensões da qualidade do ensino, tais como a interação em grupo e relacionamento individual.
 6. **nível do curso** - Aleamoni e Hexner (1980, *apud* Aleamoni, 1998) analisaram oito investigações que tratam do relacionamento “nível do curso” e “avaliações dos alunos” e reportam que esses estudos não encontraram relacionamentos significativos entre o nível do curso (1º ano, 2º ano, 3º ano etc.) e as avaliações dos professores pelos alunos. Frey *et al.* (1975) relata que as dimensões do instrumento *ENDEAVOR*, exceto a dimensão “carga de trabalho”, estão correlacionadas positivamente e significativamente com os semestres de estudo dos alunos. Donaldson *et al.* (1993), Conran (1991), Goldberg e Callahan (1991), Moristsch e Sutter (1988), (*apud* ALEAMONI, 1999) encontraram correlação entre o nível do curso e as avaliações. Capelleras e Veciana (2000), relatam em sua pesquisa, que os alunos do 1º ciclo, em média, valoram as cinco dimensões de seu instrumento mais favoravelmente que os alunos do segundo ciclo.
 7. **classificação do professor** - Segundo Marsh e Roche (1997), todas as pesquisas que tratam do assunto reportam correlações mistas entre a posição do docente (graduado, mestrado, doutorado, pós-doutorado) e a avaliação estudantil, as quais indicam pouco ou nenhum efeito.
 8. **gênero do professor** - Pesquisas visando correlacionar as avaliações dos alunos com o sexo do professor, revelam resultados contraditórios e não existe significância estatística

suficiente para concluir que os professores homens são melhores avaliados que as mulheres ou vice-versa.

9. **disciplina acadêmica** - Feldman (1978, *apud* CASHIN, 1995), revisou algumas pesquisas, evidenciando que as disciplinas das áreas humanas e artes tendem a ser avaliadas pelos discentes mais favoravelmente do que as disciplinas das áreas sociais, as quais por sua vez receberam avaliações maiores do que as disciplinas das áreas de matemática, física ou engenharia. Marsh e Roche (1997), argumentam que são necessárias pesquisas adicionais nessa área, para tornar o assunto mais claro.

10. **condições administrativas**. Braskamp e Ory (1994), Centra (1993), Feldman (1979), Marsh e Dunkin (1992), (*apud* CASHIN, 1995) reportam que os alunos tendem a majorar as avaliações dos docentes quando:

- a) é solicitado aos discentes que assinem o instrumento de pesquisa;
- b) o professor esteja presente no momento da aplicação do instrumento de pesquisa; e
- c) as informações decorrentes das análises das avaliações não forem utilizadas exclusivamente para melhoria da qualidade do ensino.

Segundo Greenwald e Gillmore (1997), existem cinco teorias que justificam o relacionamento entre as notas médias das classes com as avaliações médias da classe de alunos: 1) a correlação positiva entre nota e avaliação é decorrente da eficácia do ensino do professor. Nesse caso, a qualidade instrucional do professor é uma terceira variável que explica a correlação entre notas e avaliações, logo as notas não representam viés sobre as avaliações; 2) a motivação acadêmica geral dos alunos influencia positivamente ambos, notas e avaliações, isso porque cursos que atraem alunos altamente motivados faz com que eles se esforcem bastante e aprendam muito e, conseqüentemente, avaliam melhores os professores. A motivação do aluno tem sido sugerida como a terceira variável que justifica o relacionamento nota/avaliação e representa uma influência legítima sobre as avaliações; 3) a

motivação dos alunos, específica ao curso ou à disciplina, influencia ambos, notas e avaliação. Classes com alunos interessados pela disciplina tendem a avaliar favoravelmente o professor, embora não fique claro que o interesse do aluno existiu antes do início da disciplina ou se foi despertado pelo professor. Essa teoria retém a interpretação de que as avaliações medem a eficácia do ensino; 4) alunos inferem a qualidade do ensino do professor a partir das notas recebidas. Nesse caso, os bons alunos supostamente aprendem mais e avaliam melhor o professor, enquanto que os alunos sem potencial aprendem pouco e não reconhecem o esforço do professor. Desse modo, a nota pode vir a se constituir um viés nas avaliações dos alunos sem potencial e que não reconhecem o bom professor; 5) alunos atribuem avaliações elevadas a seus professores em decorrência de brandura nas notas. Nesse caso, as avaliações são enviesadas, dado que a nota atribuída pelo professor ao aluno não corresponde a uma aprendizagem real. Pesquisas nessa área aparecem mais no contexto de averiguar que a brandura pode representar apenas influências pequenas e ignoráveis nas avaliações dos alunos.

De acordo com Mckeachie (1979), embora muitas pesquisas³ argumentem que existam muitas variáveis que provocam algum viés nas avaliações dos alunos, elas revelam que estas variáveis supostas “enviesantes” tendem a mostrar pouco ou nenhum relacionamento com as avaliações dos professores pelos alunos. Entre as principais variáveis do aluno que podem estar relacionados com as suas avaliações destaca-se a motivação do discente e notas esperadas ou reais. Quanto as variáveis do curso, talvez a disciplina acadêmica necessite de controle. Quanto a variável carga trabalho/dificuldades, essas relacionam-se positivamente com as avaliações dos alunos, ou seja, os alunos avaliam positivamente as disciplinas em que eles necessitam estudar bastante.

³ É preciso deixar claro que as pesquisas relacionando o tipo de disciplina com as avaliações dos docentes pelos discentes ocorreram em universidades americanas e que pesquisas em universidades brasileiras devem ser feitas.

Segundo Mckeachie (1997), as universidades não deveriam dar muita ênfase às possíveis variáveis enviesantes, já que o efeito que essas exercem nas avaliações dos alunos é considerado por muitos pesquisadores irrelevantes. Para ele, as universidades, através de seus órgãos competentes, deveriam assegurar um uso mais efetivo para os relatórios das avaliações dos docentes pelos discentes, ou seja, assegurar a validade de uso.

5 AVALIAÇÃO DAS PROPRIEDADES DO INSTRUMENTO DE AVALIAÇÃO MEDIANTE MODELOS DE ANÁLISE FATORIAL EXPLORATÓRIA

5.1 Introdução

As propriedades estatísticas do instrumento de avaliação, ou seja, as propriedades de confiabilidade e de validade das diversas dimensões que, *a priori*, são operacionalizadas no instrumento de avaliação da docência, em um primeiro momento, podem ser investigadas pelo uso das técnicas de análise fatorial exploratória. Desse procedimento, resultam hipóteses que devem ser contrastadas a seguir, através dos modelos de análise fatorial confirmatória.

Análise fatorial é uma técnica multivariada, usada para analisar os interrelacionamentos 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 intercorrelacionamentos existentes entre os indicadores. Supõe-se que um conjunto de indicadores, pertencente ao domínio de indicadores, seja intercorrelacionado devido ao efeito de um fator comum que produza essa associação.

Em um modelo de análise fatorial exploratória, a variabilidade de qualquer indicador (X_i), é explicado por duas fontes de variação, quais sejam: a variabilidade do fator comum (ξ) e a variabilidade da parte específica (U_i). A variância do fator comum representa o que os indicadores intercorrelacionados do instrumento têm em comum e é expressa pelas cargas fatoriais (λ_i) no fator latente (ξ), e é essa que constitui a questão de validade dos indicadores; ou seja, quanto do fator comum (ξ) é representado empiricamente pelos indicadores associados ao fator comum. O restante da variância dos indicadores não contribui com a validade do construto ou fator comum, sendo representado pela sua unicidade (U_i), que representa tanto o que é específico de cada uma delas quanto os erros de medida.

Considerando que a estrutura de um questionário seja interpretada pela existência de dois fatores comuns, o modelo de análise fatorial exploratória, com \mathbf{K} indicadores, terá a seguinte representação, conforme diagrama de caminhos, representado pelo diagrama 1.

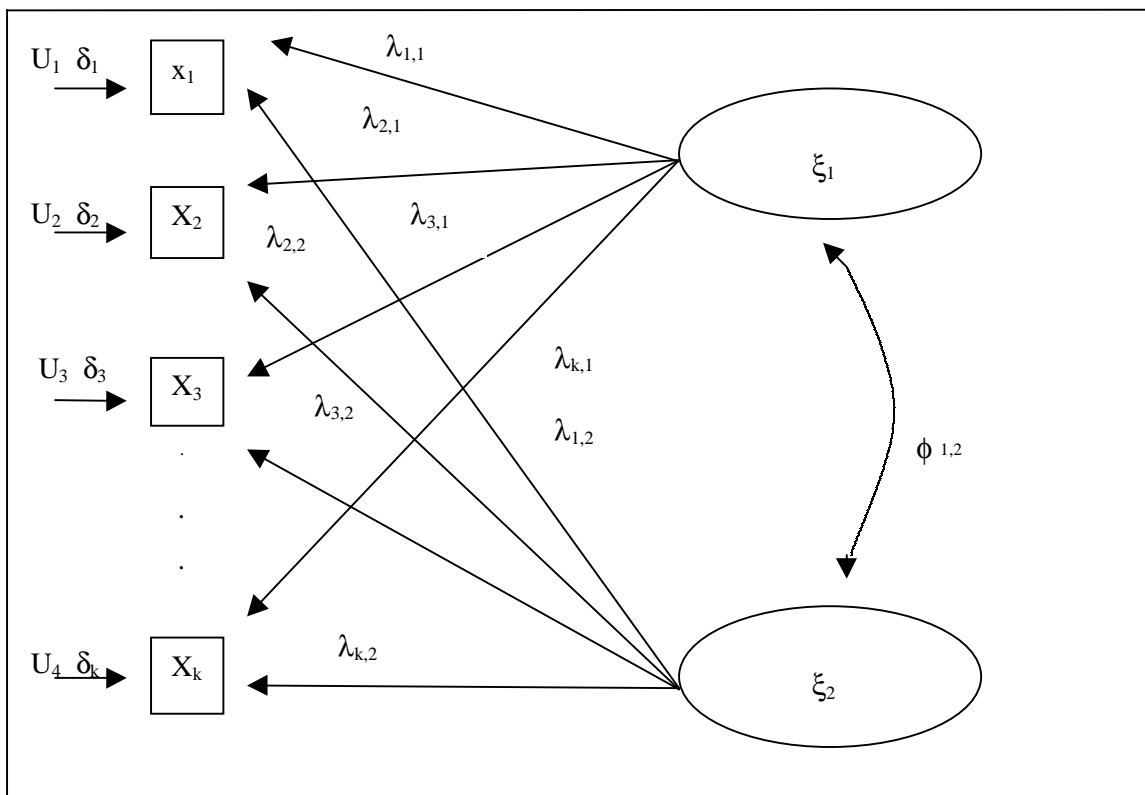


Diagrama 1: Representação do modelo fatorial com \mathbf{K} indicadores e dois fatores comuns

Na representação do modelo de análise fatorial do diagrama 1, \mathbf{X}_i ($i = 1, 2, 3, \dots, k$) são os indicadores; ξ_j ($j = 1, 2$) são os construtos que explicam os interrelacionamentos entre os indicadores; U_i ($i = 1, 2, 3, \dots, k$) corresponde a parte específica de cada indicador, \mathbf{X}_i ; λ_{ij} associada a cada flecha são as saturações, que representam a relação entre o fator comum e o indicador; δ_i é o efeito (correlação) da variabilidade associada à parte específica sobre os indicadores.

Considerando que a estrutura de um instrumento avaliativo seja interpretada pela existência de r fatores comuns ($r \leq k$) e \mathbf{K} indicadores, o modelo de análise fatorial geral que

expressa a dependência entre o indicador (X_i) e todos os fatores comuns (ξ_j) e específicos (U_i) é definido por:

$$X_i = \lambda_{i1}\xi_1 + \lambda_{i2}\xi_2 + \lambda_{i3}\xi_3 + \dots + \lambda_{ir}\xi_r + \delta_i U \quad (14)$$

Em que:

X_i : corresponde aos indicadores do instrumento, $i=1, 2, \dots, k$;

$\lambda_{i,j}$: são as cargas (correlações) dos indicadores no fator comum, $i = 1, 2, \dots, k$ e $j = 1, \dots, r$;

ξ_j : corresponde aos fatores comuns, $j = 1, 2, \dots, r$;

δ_i : corresponde à carga (correlação) da especificidade de X_i ; e

U_i : parte específica de X_i .

O modelo (14), na sua forma matricial, assume a seguinte forma:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \dots \\ X_k \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \dots & \lambda_{1r} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \dots & \lambda_{2r} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \dots & \lambda_{3r} \\ \dots & \dots & \dots & \dots \\ \lambda_{k1} & \lambda_{k2} & \lambda_{k3} \dots & \lambda_{kr} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \dots \\ \xi_r \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_k \end{bmatrix}$$

que na sua forma simplificada é equivalente a:

$$X = \Lambda \xi + U \quad (15)$$

Em que Λ corresponde a matriz das cargas fatoriais (λ_i).

As seguintes hipóteses devem ser levadas em consideração, durante o processo de estimação dos coeficientes λ_{ir} , δ_i do modelo (15).

- 1) Os \mathbf{k} indicadores são padronizados. Logo, cada um deles possui média zero e variância igual à unidade. Quando se trabalha com a matriz de correlação, os elementos da diagonal principal é composta por elementos iguais a 1. Ou seja:

$$\mathbf{X} \rightarrow N_{\mathbf{k}}(0, \Sigma) \quad (16)$$

- 2) Os \mathbf{r} ($\mathbf{r} \leq \mathbf{k}$) fatores comuns possuem cada um deles média igual a zero e variância igual a um. A matriz de covariância é a matriz identidade, logo são fatores intercorrelacionados, ou seja:

$$\xi \rightarrow N_{\mathbf{r}}(0, \mathbf{I}) \quad (17)$$

- 3) Os \mathbf{K} fatores específicos possuem cada um deles média igual a zero, estão incorrelacionados entre si e sua matriz de covariância (Ω) é diagonal. Ou seja:

$$\mathbf{U} \rightarrow N_{\mathbf{K}}(0, \Omega) \quad (18)$$

- 4) Os fatores comuns ξ_j e as variáveis específicas, U_i , estão incorrelacionadas.

O problema matemático da análise fatorial consiste em decompor a variabilidade dos indicadores, \mathbf{X}_i , matriz Σ , em duas partes: uma relacionada com o que é comum, matriz Λ , e outra com que específico, matriz Ω . De modo que:

$$\Sigma = \Lambda \Lambda' + \Omega \quad (19)$$

Explicitando o modelo (19) em forma matricial tem-se:

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13\dots} & \rho_{1i} \\ \rho_{21} & 1 & \rho_{23\dots} & \rho_{2i} \\ & & \ddots & \vdots \\ \rho_{k1} & \rho_{k2} & \rho_{k3\dots} & 1 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13\dots} & \lambda_{1i} \\ \lambda_{21} & \lambda_{22} & \lambda_{23\dots} & \lambda_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{k1} & \lambda_{k2} & \lambda_{k3} & \lambda_{ki} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \lambda_{21} & \dots & \lambda_{k1} \\ \lambda_{12} & \lambda_{22} & \dots & \lambda_{k2} \\ \lambda_{13} & \lambda_{23} & \dots & \lambda_{k3} \\ \lambda_{1i} & \lambda_{2i} & \ddots & \lambda_{ki} \end{bmatrix} + \begin{bmatrix} \delta_1^2 & 0 & 0\dots & 0 \\ 0 & \delta_2^2 & 0\dots & 0 \\ 0 & 0 & \delta_3^2\dots & 0 \\ 0 & 0 & 0\dots & \delta_k^2 \end{bmatrix}$$

A variância de cada uma dos indicadores apresentados do modelo (22), X_i ($i = 1, 2, \dots, K$) é definida por:

$$\begin{aligned} \text{VAR}(X_i) &= \text{VAR}(\lambda_{i1}\xi_1 + \lambda_{i2}\xi_2 + \lambda_{i3}\xi_3 + \dots + \lambda_{ir}\xi_r + \delta_i U_i) \\ \text{VAR}(X_i) &= \lambda_{i1}^2 \text{VAR}(\xi_1) + \lambda_{i2}^2 \text{VAR}(\xi_2) + \lambda_{i3}^2 \text{VAR}(\xi_3) + \dots + \lambda_{ir}^2 \text{VAR}(\xi_r) + \delta_i^2 \text{VAR}(U_i). \end{aligned}$$

Levando em consideração que os fatores comuns e específicos são padronizados, a variância de (X_i) torna-se:

$$\text{VAR}(X_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \lambda_{i3}^2 + \dots + \lambda_{ir}^2 + \delta_i^2 \quad (20)$$

ou alternativamente, tem-se que :

$$1 = \lambda_{i1}^2 + \lambda_{i2}^2 + \lambda_{i3}^2 + \dots + \lambda_{ir}^2 + \delta_i^2 \quad (21)$$

A soma das r primeiras parcelas de (21) denomina-se “comunalidade” (h^2) e indica a medida de quanto da variância de um indicador é explicado pelos fatores decorrentes da análise fatorial (PEREIRA, 1999). Como δ_i^2 é variabilidade da especificidade do indicador, que é decorrente das fontes sistemáticas e aleatórios de erro, a variância do indicador X_i é definido em termos das seguintes variâncias: fator comum, específica e do erro aleatório. Assim:

$$\text{VAR}(X_i) = h^2 + \delta_i^2 + \varepsilon_i^2 \quad (22)$$

ou

$$1 = h^2 + \delta_i^2 + \varepsilon_i^2 \quad (23)$$

Segundo Hair *et al.* (1999), a decomposição da variância total do indicador (X_i) em termos de sua comunalidade, de sua especificidade e o termo de erro aleatório, possibilita

perceber a diferença que existe entre a confiabilidade e a validade da medida. A confiabilidade é constituída pela variância verdadeira dos indicadores (\mathbf{X}_i), ou seja, pela variância comum e pela variância específica dos indicadores. Enquanto a validade é constituída somente pela variância comum, ou seja, a covariância entre o indicador (\mathbf{X}_i) e os fatores comuns ou latentes. Assim, a confiabilidade (r) e a validade (θ) das medidas são respectivamente iguais a:

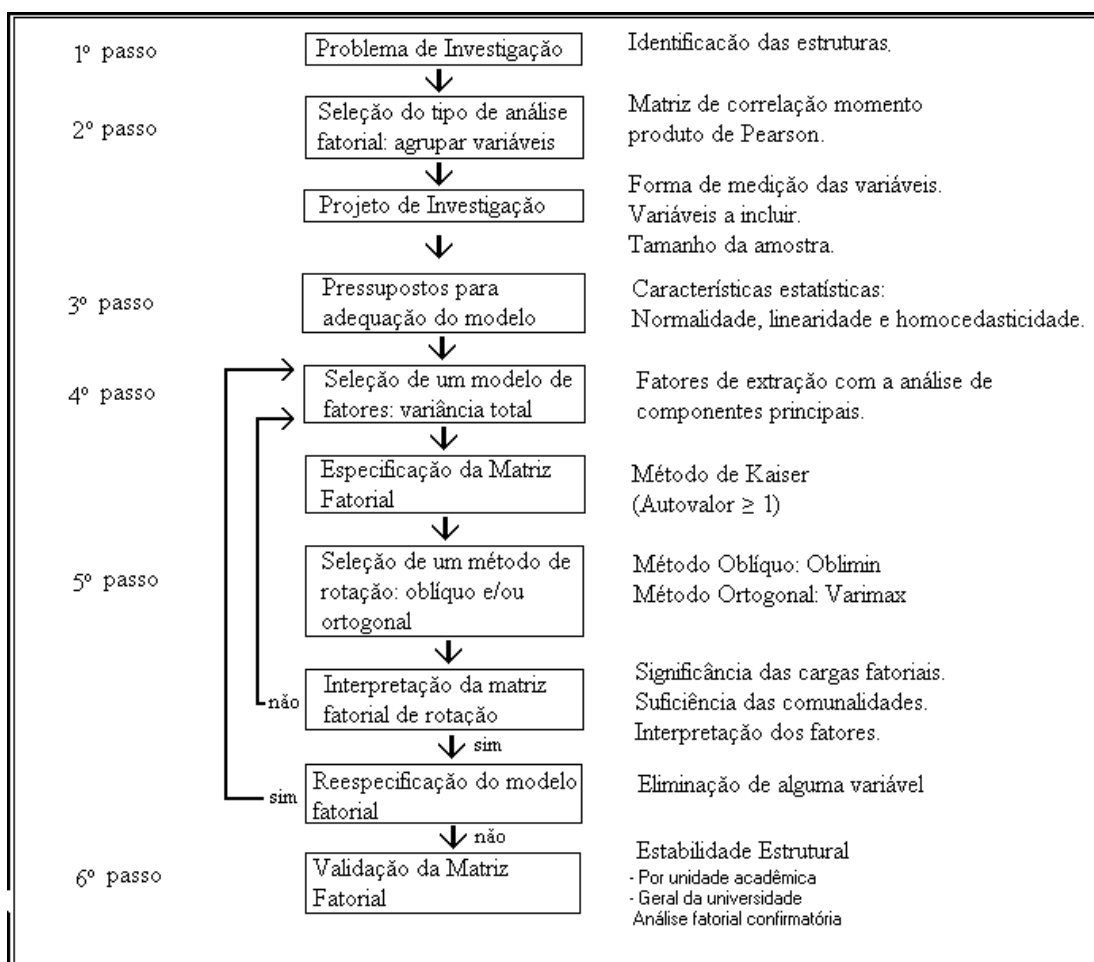
$$r = 1 - \varepsilon_i^2 \quad (24)$$

$$\theta = h^2 \quad (25)$$

O modelo (24) evidencia que, quando não existir o erro aleatório no processo de medição das indicadores (\mathbf{X}_i), a confiabilidade será igual a um. A comunalidade (h^2) representa o que os indicadores do teste têm em comum, sendo manifestada pelas cargas fatoriais nos fatores comuns (ξ). É essa que constitui a questão da validade do teste. Quanto maior for a grandeza da carga fatorial do indicador (\mathbf{X}_i) em um único fator comum (ξ), maior o poder de discriminação do indicador. Isso significa que as variabilidades dos indicadores entram com toda a sua magnitude na especificação de um único fator comum. Um fator ou construto é válido quando quão bem os indicadores representam um único construto hipotético, ou seja, as cargas fatoriais, que indicam que a covariância entre o construto e os indicadores são elevadas. Assim, estando o construto consistentemente determinado, é possível interpretar o seu significado.

5.2 O processo de análise fatorial

O fluxograma 3, adaptado de Hair et al (1999), apresenta os passos a serem utilizados na determinação da estrutura fatorial, que tem como propósito validar os distintos conceitos implícitos no instrumento de avaliação da docência pelos alunos das instituições de ensino superior



Fluxograma 3: Passos utilizados na extração dos fatores do modelo de análise fatorial exploratória

5.2.1 Problema de investigação

O objetivo geral do uso das técnicas analíticas de análise fatorial exploratória é o de resumir as informações contidas em uma série de variáveis originais ou indicadores, em uma série menor de dimensões compostas ou fatores teóricos novos com perda mínima de informação, ou seja, determinar uma estrutura subjacente aos dados que conservem o máximo possível da variabilidade manifestadas pelos distintos indicadores do instrumento de avaliação. Segundo Hair *et al.* (1999), é possível pelo uso das técnicas de análise fatorial definir as construções teóricas ou dimensões que se supõe servirem de base para as variáveis originais ou indicadores.

Pretende-se com essa metodologia de análise multivariada confirmar através dos níveis de satisfação dos alunos, as dimensões da qualidade do ensino projetadas para medir a qualidade do ensino no âmbito da sala de aula das instituições de ensino superior.

5.2.2 Seleção do tipo de análise fatorial

O objetivo da investigação é confirmar as oito dimensões da qualidade do ensino do instrumento “Avaliação pelos Alunos do Desempenho na Disciplina”. Portanto, espera-se que os 29 indicadores avaliados pelos alunos estejam cada um deles altamente saturados com o conceito que foi designado para medir.

Para identificar os fatores que formam a base dos indicadores, é necessário investigar os interrelacionamentos dessas correlações por meio do coeficiente de correlação linear de Pearson (r), que irá indicar a magnitude das correlações entre pares de indicadores. Todas as correlações são disponibilizadas para análises dos fatores através da matriz de correlação ou variância-covariância.

Uma matriz de correlação representa os relacionamentos entre o conjunto de indicadores investigados. Nessa matriz, os valores localizados na diagonal principal serão iguais à unidade, isso porque cada uma das variáveis se correlacionará perfeitamente consigo mesma. Os elementos das diagonais secundárias são as correlações entre todos os pares de indicadores. A matriz de variância-covariância apresenta as covariações entre os pares de indicadores. Nele, os valores localizados na diagonal principal são um, isso porque cada uma das variáveis é padronizada, portanto a sua variância é um. Os elementos das diagonais secundárias são as covariâncias entre todos os pares de indicadores.

Hair *et al.* (1999), apresenta dois tipos de análise fatorial: análise fatorial do tipo R e do tipo Q. A análise do tipo R faz o uso da matriz de correlações ou de variância-covariância com o propósito de identificar as dimensões que são latentes, quais sejam, aquelas não facilmente observadas. A análise fatorial do tipo Q preocupa-se em agrupar indivíduos em grupos ou *cluster*. Do exposto, concluiu-se que a análise fatorial tipo R, qual faz o uso da matriz de correlação ou variância-covariância, é o tipo de análise fatorial adequada para identificar a estrutura fatorial subjacente aos dados.

5.2.3 Projeto de investigação

Nessa etapa, Hair *et al.* (1999) chama a atenção para os seguintes aspectos: números de indicadores, forma de medição e o tamanho amostral que possibilite a obtenção de um modelo simplificado que preserve o máximo possível das variabilidades dos indicadores por meio de uma técnica de estimação. O número de indicadores intercorrelacionados para representar um conceito é cinco ou mais; uma única indicadora tem pouca utilidade.

Proceder a análise fatorial com a média aritmética de cada um dos indicadores, permitirá trabalhar com uma medida contínua pura, ou seja, medida isenta de erros. O uso adequado da análise fatorial decorre do fato dos indicadores serem contínuas (HAIR *et al.*, 1999). O formulário “Avaliação pelos Alunos do Desempenho na Disciplina” possui 35 indicadores; considerando que a unidade de análise é a média desses indicadores, é necessária a seleção aleatória de 175 disciplinas para o uso adequado dos modelos analíticos de análise fatorial. Ao utilizar esse procedimento de amostragem, as cargas fatoriais superiores a 0,40 alcançam significância estatística ao nível de significância de 5% (HAIR *et al.*, 1999).

5.2.4 Pressupostos para a adequação do modelo

Para Hair *et al.* (1999), o sucesso das conclusões decorrentes da submissão dos dados aos modelos analíticos de análise fatorial depende das características estatísticas de normalidade, linearidade e homocedasticidade dos indicadores projetados para contribuir na operacionalização de determinados conceitos.

Quanto às distribuições dos distintos indicadores, eles devem exibir um erro padrão de simetria e curtose inferior a $\pm 1,96$ para que a distribuição dos indicadores seja considerada normal, ao nível de significância de 5% (HAIR *et al.*, 1999). A suposição da normalidade dos dados se faz necessária, para se testar as hipóteses colocadas à prova por meio dos modelos de análise fatorial confirmatório, utilizado para confirmar a natureza dos construtos teóricos resultantes do processo de análise fatorial exploratória. Os modelos confirmatórios estimados exibem para a sua *performance*, o valor de *t* de *student* para cada um dos indicadores associados ao construto, além da estatística χ^2 (Qui-quadrado) e outros indicadores da qualidade do ajuste, os quais segundo Hair *et al.* (1999), são enfraquecidos pela falta de normalidade dos indicadores.

Quanto à homocedasticidade dos indicadores, ou seja, investigar se esses exibem variância constante, uma solução é efetuar a comparação de igualdade das matrizes de variância/covariância pelo teste *de Levene* (Hair *et al.*, 1999). Um método alternativo consiste na análise dos gráficos de dispersão, resultante do cruzamento de dois indicadores.

No que diz respeito à linearidade, observar atentamente a matriz de correlação dos indicadores, possibilita concluir se os dados são adequados para o processamento da análise fatorial. Segundo Hair *et al.* (1999), se não existir uma quantidade substancial de correlações superiores a 0,30, o uso da análise fatorial é inapropriado. Um teste estatístico, muito utilizado para testar a hipótese de que todos os elementos da diagonal principal da matriz de correlação são iguais a unidade (hipótese nula a ser testada) e os demais elementos são zero (hipótese alternativa a ser testada), é o teste de esfericidade de *Bartlett*, que é uma transformação da prova do Qui-Quadrado (χ^2). Sob a hipótese de que as distribuições dos indicadores são provenientes de uma população normal multivariada, a estatística χ^2 é definida, segundo Martínez (2000) por:

$$\chi^2_{[0,5(k^2 - k)]} = - [n - 1 - 1/6 (2k + 5)] \ln | R | \quad (26)$$

No modelo (26), n é o tamanho da amostra; K corresponde ao número de indicadores do instrumento; $|R|$ é o determinante da matriz de correlação. Aceitando-se a hipótese alternativa a partir de um nível de significância especificado, pode-se efetuar o processo de análise fatorial.

Um índice alternativo ao teste de esfericidade de *Bartlett* é o índice de *Kaiser-Meyer-Olkin* (*KMO*). Esse índice compara a grandeza dos coeficientes de correlação observados com os coeficientes de correlações parciais para o conjunto de indicadores. Segundo Martínez (2000), esse índice é definido por:

$$KMO = \frac{\sum_{i=1}^k \sum_{j \neq i}^k r_{ij}^2}{\sum_{i=1}^k \sum_{j \neq i}^k r_{ij}^2 + \sum_{i=1}^k \sum_{j \neq i}^k a_{ij}^2} \quad (27)$$

em que:

r_{ij}^2 : coeficiente de correlação simples entre os indicadores; e

a_{ij}^2 : coeficiente de correlação parcial entre os indicadores.

Quando as correlações parciais forem muito pequenas, o índice *KMO* alcança valor próximo da unidade e indicará ótima adequação dos dados para a análise fatorial. Seus valores críticos, segundo Hair *et al.* (1999) são:

- valores próximo a 0,90, indicam ótima adequação dos dados à análise fatorial;
- valores próximos a 0,80 indicam boa adequação dos dados à análise fatorial;
- valores próximos a 0,70 indicam razoável adequação dos dados à análise fatorial;
- valores próximos a 0,60 indicam pobre adequação dos dados à análise fatorial; e
- valores próximos a 0,50 ou menores indicam que os dados não são adequados para se proceder a uma análise fatorial.

Também pode-se observar a matriz de correlação parcial e a matriz de correlação antiimagem. Quando os coeficientes de correlação parcial (correlação de um indicador contra outro, controlando os efeitos de todos os outros considerados no estudo) forem pequenos, é um indicativo de que se pode explicar a variabilidade das variáveis, mediante os fatores teóricos. Se as correlações parciais forem elevadas, então não existem fatores latentes subjacentes “verdadeiros” e o emprego da análise fatorial com o propósito de obter uma estrutura reduzida é inadequado (HAIR *et al.* 1999).

A matriz de correlação antiimagem é simplesmente o valor negativo dos coeficientes de correlação parcial, cujos valores em termos absolutos devem ser pequenos para se realizar uma análise fatorial (HAIR *et al.* 1999).

5.2.5 Seleção de um modelo de fatores

Como a análise fatorial, nesse estudo, tem como objetivo encontrar uma estrutura simplificada que explique o máximo da variabilidade total dos indicadores selecionados, o método dos componentes principais é o adequado para a estimação dos fatores (HAIR *et al.* 1999). A primeira componente pode ser considerada como o melhor resumo das relações lineares existentes entre os dados. A segunda será o segundo melhor resumo com a condição de que seja ortogonal no primeiro componente principal, e assim sucessivamente.

5.2.6 Especificação da matriz fatorial

Nessa etapa, determina-se o número de fatores a ser retido pela resolução fatorial. A preocupação central é a de investigar se um número menor de fatores latentes pode explicar a covariância do conjunto de indicadores medidos com pouca perda de informação. Hair *et al.* (1999), cita que existem vários métodos que podem ser usados para decidir quantos fatores reter, quais sejam: critério da raiz latente, critério da porcentagem da variância, critério de contraste de caída e outros.

Pesquisadores, tais como Capelleras e Veciana (2000), Tejedor *et al.* (1988) e Moreira (1986) utilizaram o critério da raiz latente nos seus estudos de multidimensionalidade de seus instrumentos. Segundo Hair *et al.* (1999), a racionalidade que se usa para o critério da raiz latente é que fator latente individual deveria justificar a variância de pelo menos uma única variável. Como os indicadores são reduzidos, cada um deles tem variância igual a um (1), assim, o mais usual é considerar fatores com autovalores maiores do que um (KAISER, 1974). Outro argumento a favor do critério de Kaiser (1974) é fornecido por Hair *et al.*

(1999), justificando que o uso do *eigenvalor* estabelece um corte mais confiável quando o número de indicadores está entre 20 e 50.

5.2.7 Seleção de um método de rotação

A interpretação coerente da multidimensionalidade do instrumento é feita através das cargas fatoriais, informando que indicadores estão melhores correlacionados com as dimensões latentes ou construtos teóricos.

A derivação das dimensões latentes pelo método das componentes principais ocorre por uma seqüência de rotações de eixos que melhor expressa a dispersão dos dados. No modelo fatorial final, as variações dos indicadores estão maximizadas, enquanto as relações entre as dimensões latentes estão suavizadas. Assim, o pesquisador deve preferir interpretar os fatores latentes após a rotação da matriz fatorial. Hair *et al.* (1999) apresenta diferentes estratégias de rotação para o modelo fatorial, as quais podem agrupar-se em dois blocos:

- 1) os que mantêm a ortogonalidade dos fatores latentes – rotação ortogonal; e
- 2) os que não mantêm a ortogonalidade dos fatores latentes – rotação oblíqua.

Existem diversas estratégias de rotação ortogonal, sendo o método de rotação **VARIMAX** o mais utilizado no processo de validação de construto. O seu uso é justificado pelo fato de que o método de rotação busca minimizar o número de indicadores com altas cargas num fator latente. Segundo Kaiser (1974), o padrão fatorial obtido mediante a rotação ortogonal **VARIMAX** tende a ser mais robusto que o obtido pelo método **QUARTIMAX**, quando se analisam diferentes subconjuntos de variáveis intercorrelacionadas.

Os métodos de rotação oblíqua permitem a existência de fatores latentes correlacionados em lugar de manter a independência entre os fatores latentes rodados, dados que eles rodam livremente até alcançar a melhor descrição da estrutura dos indicadores. É

particularmente usado durante as etapas de validação de construto de instrumentos projetados para medir a qualidade educacional, obtendo, assim, construtos teoricamente significativos (TEJEDOR *et al.* 1988, MARSH, 1984). Do processo de rotação oblíqua, resulta a matriz de correlação dos fatores latentes. Quanto maior a correlação entre fatores, menor a capacidade de discriminação dos indicadores destinados a medir determinado conceito teórico.

5.2.8 Interpretação da matriz fatorial de rotação

A análise fatorial é puramente uma técnica estatística indicando qual e a que ponto os indicadores se relacionam para formar um fator latente subjacente. O significado substantivo atribuído a um fator latente está tipicamente baseado no exame cuidadoso do pesquisador do que os indicadores com cargas fatoriais significativas medem. Ciente de que essas variáveis têm em comum, é possível atribuir um significado para a especificidade de cada fator latente.

5.2.9 Reespecificação do modelo fatorial

Quando todos os indicadores manifestarem carga fatorial significativa, no seu respectivo fator latente, que teoricamente deveria estar correlacionada, a interpretação do significado de cada fator latente é feita com maior segurança. Não obstante algum indicador venha evidenciar carga fatorial significativa em dois ou mais fatores latentes, o trabalho de interpretar o significado destes fatores é muito mais complicado, uma vez que o indicador não consegue discriminar entre dois fatores latentes (HAIR *et al.* 1999).

A observação atenta da comunalidade do indicador, ou seja, a investigação da variância dessa na solução da estrutura fatorial final, flexibiliza o pesquisador a avaliar se o

indicador alcança níveis aceitáveis de explicação. Os indicadores que manifestam baixas comunalidades são indicadores livres, devendo ser removidos da análise. Optando por esse procedimento, é prudente a estimação de um novo modelo de fatores, com maior capacidade explicativa da variância total da matriz de correlação.

5.2.10 Validação da matriz fatorial

Nessa etapa, deve-se certificar de que a estrutura fatorial resultante apresenta elevado grau de generalidade para uma população. HAIR *et al.* (1999) diz que o método mais direto de validação dos resultados consiste em adotar uma perspectiva de confirmação, valorando a replicabilidade dos resultados.

A divisão da amostra em duas partes, em que uma delas se destina à análise fatorial exploratória, enquanto a outra se destina à análise fatorial confirmatória através dos modelos de equação estrutural com variáveis latentes e erro de medida, é uma alternativa recomendada para confirmar os construtos teóricos decorrentes da análise fatorial exploratória.

Tendo evidenciado a replicabilidade dos fatores latentes através dos métodos analíticos distintos, há algumas evidências de que o construto está se estabelecendo, e desse modo, atribuir um certo grau de validade para as distintas dimensões do instrumento. Destaca-se ainda a necessidade de submeter o instrumento a contextos educacionais distintos para confirmar o seu grau de generalidade.

5.2.11 Análise fatorial exploratória *versus* confirmatória

A análise exploratória ajuda o pesquisador a identificar a estrutura fatorial latente para um conjunto de indicadores, corroborando para o melhor esclarecimento de um conceito ou conceitos, não operacionalmente definido. Em contraste, a análise fatorial confirmatória é baseada numa teoria consolidada que permite ao pesquisador especificar antecipadamente um modelo. Esse modelo especifica quais indicadores carregarão em quais fatores, como também os índices de correlação entre os fatores latentes (STAPLETON, 1997).

Stevens (1996, *apud* STAPLETON, 1997), apresenta um quadro comparativo, quadro 3, chamando a atenção para as diferenças existentes entre os modelos de análise fatorial exploratória e confirmatória.

Análise exploratória: geração de teoria	Análise confirmatória: teste da teoria
<ul style="list-style-type: none"> - Determina o número de fatores latentes. - Determina se os fatores latentes estão correlacionados ou não correlacionados. - Indicadores livres de cargas se manifestam em todos os fatores latentes. 	<ul style="list-style-type: none"> - Número de fatores latentes é fixado, <i>a priori</i>. - Os fatores fixados <i>a priori</i> são tidos como correlacionados ou não correlacionados. - Os indicadores são fixadas para saturarem num fator.

Quadro 3: Quadro Comparativo entre Análise Fatorial Exploratória x Confirmatória.

5.2.12 O uso da técnica de análise fatorial exploratória para avaliar a validade de construto – algumas restrições.

Até o final dos anos oitenta, pesquisas empíricas sobre as dimensões da qualidade do ensino eram realizadas, utilizando-se das técnicas de análise fatorial exploratória. Quando um conjunto de indicadores saturava significativamente um fator, era o suficiente para dotar o instrumento de avaliação de um certo grau de validade de construto (WHITELY e DOYLE, 1978; UMBLE e WHITTEN, 1977; COVERT e MASON, 1974; SHERMAN e BLACKBURN, 1975; FRENCH-LAZOVIK, 1974).

Instrumentos que tiveram a sua multidimensionalidade testada pelo procedimento analítico de análise fatorial exploratória, tais como o de Blanco e Paz (1988), Tejedor *et al.* (1988), Frey *et al.* (1975), apresentaram um grande componente principal que explica mais que 30% da variabilidade das respostas dos alunos. Em geral, a maioria dos estudos de fator pode ser interpretada por providenciarem evidência para um componente global ao invés de vários fatores distintos e específicos.

Segundo D'Apollonia e Abrami (1997), a redução da matriz de dados a componentes específicos por componentes principais, não providenciam evidência incontestável de que os formulários de avaliação dos alunos medem dimensões latentes instrucionais distintas.

Mesmo que se tenha alcançado um modelo fatorial adequado, capaz de ser replicado em diferentes contextos, ainda é necessário que a estrutura fatorial obtida seja confirmada por modelos estatísticos capazes de expressar cada indicador do instrumento de avaliação como função de algum construto teórico ou variável latente, bem como a sua significância, segundo um modelo probabilístico. São essas as razões:

- a técnica de análise fatorial exploratória não assume nenhum modelo de medida para os distintos indicadores do instrumento, sua única finalidade é a de reduzir a dimensionalidade;
- a análise fatorial, em seu contexto exploratório, é geralmente considerada mais uma geradora de teoria do que um procedimento de teste de teoria;
- não existe uma cota mínima para a variância total explicada pela estrutura fatorial escolhida. Existem diferentes critérios que podem ser utilizados por diferentes pesquisadores como critério de corte. Assim, para o mesmo problema, podem existir distintas soluções fatoriais;

- não é possível determinar uma única função para cada indicador a partir de um único fator comum, ou seja, *a priori*, todos os indicadores se acham afetados por todos os fatores comuns, dificultando reconhecer o que de fato o indicador está medindo;
- a diversidade de métodos de rotação, usada arbitrariamente, conduz a interpretações distintas; e
- é pouco realista supor que, na área educacional, os fatores comuns sejam não correlacionados.

Em face do exposto, dotar determinado instrumento de avaliação de certo grau de validade de construto com base na estrutura fatorial decorrente da análise fatorial exploratória é insuficiente, pois a análise fatorial exploratória necessita de postulações rígidas que nem sempre refletem a realidade dos fatos. Por outro lado, é uma ferramenta que ajuda o pesquisador a formular ou reformular conceitos não claramente definidos. A análise fatorial confirmatória testa a hipótese da veracidade desses conceitos latentes ou construtos, ou seja, confirma ou não a legitimidade dos construtos operacionalizados no instrumento através de dados provenientes da valoração do instrumento pelos alunos.

Segundo Jöreskog (1969), o modelo de análise fatorial confirmatória corrige as deficiências próprias do modelo e conduz a uma maior certeza das hipóteses que devem ser contrastadas através de modelos que expliquem todos os intercorrelacionamentos existentes na estrutura de um questionário. Portanto, a avaliação das propriedades de validade e confiabilidade de um questionário implica em um conjunto de decisões que se sustentam em contrastes de hipóteses, corretamente formuladas.

6 AVALIAÇÃO DAS PROPRIEDADES DO INSTRUMENTO DE AVALIAÇÃO MEDIANTE MODELOS DE ANÁLISE FATORIAL CONFIRMATÓRIA

6.1 Introdução

O uso dos modelos de análise fatorial confirmatória, caso particular dos modelos de equações estruturais com variáveis latentes e erros de medidas desenvolvidos por Jöreskog (1973), Browne (1974) e Bentler (1980), constitui uma poderosa técnica multivariada cada vez mais utilizada em todas as áreas das ciências.

Segundo Hair *et al.* (1999), as razões pelas quais é utilizada em distintas áreas, deve-se a sua versatilidade em proporcionar um método direto de tratar com múltiplas relações simultaneamente com eficácia estatística, sua capacidade para avaliar as relações exaustivamente e proporcionar uma transição do exploratório ao confirmatório.

Na área educacional, os trabalhos de Marsh (1991), Ramos (1986, 1987), Val (1994), Marsh e Hocevar (1984) e Mateo e Fernández (1992) confirmam a versatilidade dos modelos de análise fatorial confirmatória no estudo da multidimensionalidade de instrumentos destinados à avaliação da docência. Esses trabalhos indicam que a partir da década dos anos 80, a abordagem de validação de construto utiliza-se da metodologia de análise fatorial confirmatória, para a confirmação da multidimensionalidade, de instrumento destinado à avaliação da docência e nas investigações de variáveis que podem envolver os níveis de satisfação dos alunos.

Segundo Ramos (1987), a análise fatorial confirmatória deve ser utilizada para provar a qualidade das distintas estruturas fatoriais alternativas que, derivadas da teoria, apresentam-se como explicativas da existência de um ou vários construtos subjacentes em um campo de covariações ou correlações entre o domínio de indicadores selecionados, através de

um modelo causal. Um modelo causal é definido pela representação formal de uma teoria ou de uma parte dela, mediante um modelo estrutural e um modelo de medida endógeno e/ou exógeno. São utilizados para modelar simultaneamente os vários indicadores medidos pelo pesquisador e as variáveis latentes, possibilitando a análise conjunta de todos os intercorrelacionamentos (GARCIA e MARTINEZ, 2000). O modelo de medida expressa a combinação linear entre os indicadores (variáveis observáveis) e os construtos teóricos previamente estabelecidos através dos modelos exploratórios.

É a forma mais idônea para a avaliação do instrumento de avaliação, dado que a maioria dos instrumentos que tratam da avaliação de professores considera que o conceito qualidade de ensino deve ser tratado como um conceito multidimensional. Esses possibilitam designar com base em um arcabouço teórico, os indicadores específicos aos construtos ou dimensões latentes, originando um modelo que pode ser estimado através de métodos estatísticos, bem como de indicadores da qualidade de ajuste absoluto, incremental e de parcimônia, que possibilite, em decorrência dos resultados, dotar o instrumento de medida de um certo grau de validade de construto. É particularmente útil para a avaliação das propriedades de validade e de confiabilidade de um instrumento de avaliação (HAIR *et al.*, 1999; RAMOS, 1997; GARCIA; MARTÍNEZ, 2000; PÉREZ; SÁNCHEZ, 1999).

O modelo de análise fatorial confirmatória difere do modelo de análise fatorial exploratória, quanto aos aspectos:

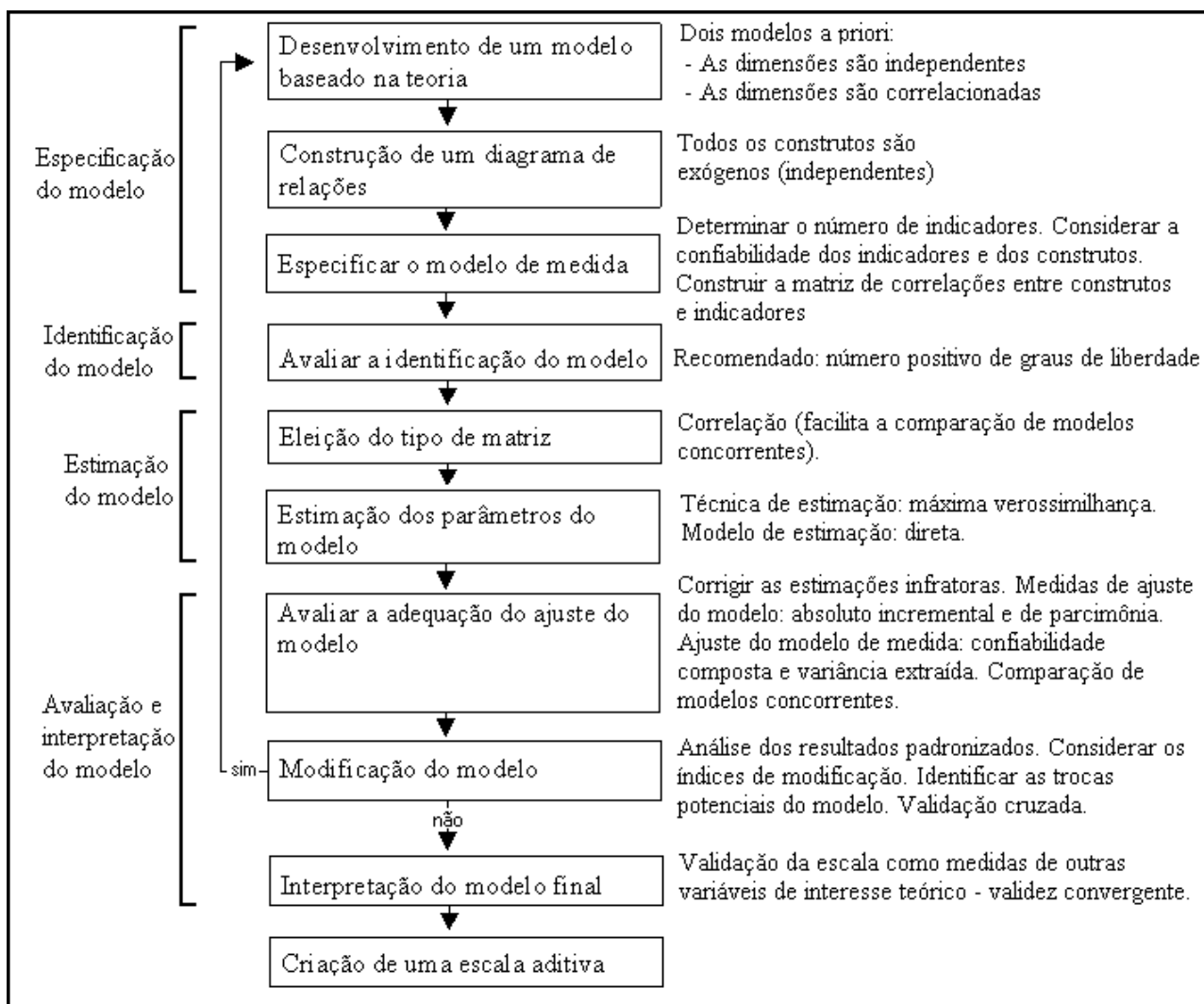
- 1) cada indicador satura em um único fator latente, restrição que torna mais fácil a interpretação dos fatores;
- 2) permite testes estatísticos para as hipóteses decorrentes da teoria, possibilitando a generalização dos resultados para a população;
- 3) possibilita o uso de medidas congênicas, ou seja, não assume a igualdade das saturações das indicadores sobre o fator comum e nem da variância do erro; e

4) permite que a parte específica do indicador esteja correlacionado com outras fontes de erro.

6.2 Fases para o desenvolvimento de um modelo de equações estruturais

Para o desenvolvimento de um modelo de equação estrutural, visando a contrastar as hipóteses decorrentes da análise exploratória, Hair *et al.* (1999), recomenda que as seguintes fases sejam levadas em consideração: a especificação do modelo, a identificação do modelo, a estimação do modelo, a avaliação e interpretação do modelo proposto.

O fluxograma 4 apresenta os passos adaptado de Hair *et al.* (1999), necessários para testar as hipóteses levantadas sobre a multidimensionalidade do instrumento de avaliação da docência pelos alunos das instituições de ensino superior, com o propósito de atribuir certo grau de validade de construto ao instrumento. Esse fluxograma propõe uma trajetória para operacionalizar as etapas 1, 2, 3 e 4, do modelo de Ramos (1986), com vistas à validação de construto.



Fluxograma 4: Passos utilizados na confirmação dos fatores decorrentes das análises exploratórias

6.3 Especificação do modelo

Nessa etapa, o pesquisador deve desenvolver um modelo baseado na teoria, construir um diagrama de relações e especificar o modelo de medida.

6.3.1 Desenvolvimento de um modelo baseado na teoria

Durante o processo de desenvolvimento do modelo, segundo Garcia e Martinez (2000), é necessário estabelecer todas as relações de dependência entre os diversos indicadores que se consideram relevantes na explicação de um determinado construto, de tal forma que a força e a certeza de tal relação depende somente e exclusivamente do apoio teórico que se tenha. A exclusão ou omissão de algum indicador relevante, reconhecido durante o processo de especificação, origina um erro de especificação, conseqüentemente, estar-se negando para a avaliação a importância de outros indicadores relevantes para o processo ensino-aprendizagem.

Tendo o pesquisador um suporte teórico bastante sólido, que lhe possibilite avaliar a força com que os indicadores operam em um modelo teórico, então a elaboração de modelos fatoriais confirmatórios com relações causais bem especificadas, será proposto para ser submetido à prova.

6.3.2 A construção de um diagrama de relações

Nessa etapa, as amostras de indicadores, associados a seus respectivos construtos, definidos, *a priori*, através de estudo lógico e/ou estudo empírico, são apresentadas por um diagrama de caminho, sobre a hipótese da existência de relacionamento linear causal entre os indicadores e os seus respectivos construtos e/ou entre os construtos teóricos.

Convencionalmente, os retângulos representam os indicadores, as elipses representam os fatores comuns, flechas unidirecionais entre fatores comuns e indicadores expressam as saturações, flechas bidirecionais indicam correlações entre fatores comuns (HAIR *et al.* 1999).

Como ilustração, o diagrama 2, mostra como se saturam treze indicadores que fazem parte do instrumento de Avaliação pelos alunos do desempenho nas disciplinas da UNISUL. Destaca-se que o bloco é composto por 29 indicadores e que o modelo completo está exposto no capítulo seguinte. Os indicadores são denominados X_i , os fatores comuns ou latentes são indicados pela letra grega ξ , os erros de medidas são indicados pela grega δ (delta), enquanto o símbolo grego ϕ indica as correlações entre os fatores comuns ou latentes.

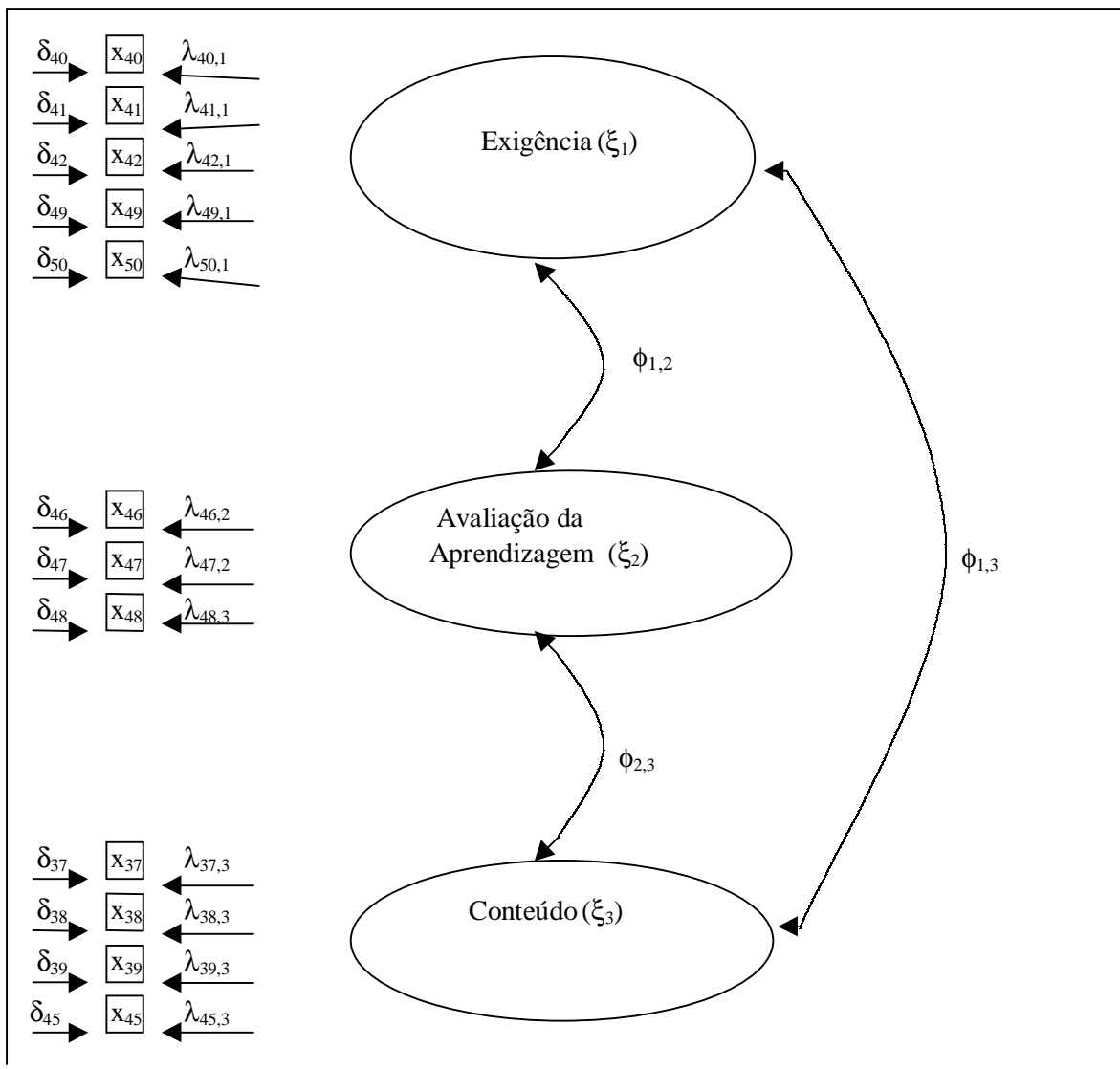


Diagrama 2: Diagrama de caminhos das características da qualidade docente (três fatores latentes correlacionados)

No diagrama 2, as flechas unidirecionais entre fator comum e indicador indicam uma influência direta do fator comum sobre o indicador. Os parâmetros (λ), associados a cada flecha, são as saturações que representam a relação entre o indicador e o fator comum. Cada saturação apresenta dois subíndices, o primeiro corresponde à identificação do indicador, em que chega a flecha (efeito), e o segundo, corresponde à identificação do fator comum de saída (causa). Todos os coeficientes λ , ϕ e δ , são, *a priori*, já estabelecidos de forma exploratória e espera-se a sua confirmação, pela submissão dos dados a uma análise fatorial confirmatória.

Os fatores comuns de um diagrama de caminho são classificados em duas classes: exógenos e endógenos. Os fatores comuns ou construtos exógenos, também denominados independentes, não são explicados por nenhuma dos indicadores do modelo, isto é, não existem flechas na direção desses construtos teóricos.

Um construto é denominado endógeno ou dependente quando são explicados por outras variáveis latentes do modelo, isto é, existem flechas apontando para a direção desses construtos. Os construtos ξ_1 , ξ_2 e ξ_3 , apresentados no quadro 9, são exógenos, dado que não existem flechas para a direção desses construtos.

6.3.2.1 Modelo estrutural

Após o desenvolvimento do modelo teórico e de sua representação por um diagrama de caminhos, deve-se especificar o modelo em termos mais formais. Esse procedimento consiste em expressar o diagrama de caminhos por um conjunto de equações lineares que definem as equações estruturais que vinculam os construtos, o modelo de medida que especifica que indicadores se associam ao construto e uma série de matrizes que indicam qualquer correlação entre os construtos.

O modelo estrutural é uma rede de relações em forma de um conjunto de equações lineares que engloba variáveis latentes endógenas e variáveis latentes exógenas. O objetivo é vincular as definições operacionais dos construtos com a teoria e, dessa forma, facilitar o contraste de hipóteses. O modelo também descreve as relações entre construtos endógenos e exógenos.

Usando a notação descrita por Jöreskog e Sörbom (1969, *apud* JOHNSON; WICHERN, 1998), a forma matricial abreviada da representação do modelo estrutural é a seguinte:

$$\eta = \beta\eta + \Gamma\xi + \zeta \quad (28)$$

em que:

η (n x 1): é o vetor das variáveis latentes endógenas;

β (n x n): é a matriz de coeficientes endógenos;

Γ (n x m): é a matriz de coeficientes exógenos;

ξ (m x 1): é o vetor de variáveis latentes exógenas; e

ζ (n x 1): é o vetor dos termos de perturbações, decorrente de perturbações aleatórias e/ou sistemáticas. Os sub-índices n e m, expressam respectivamente os números de construtos endógenos e exógenos do modelo de equação estrutural.

Os modelos de medidas são dois sub-modelos do modelo de equação estrutural, que procuram especificar os indicadores para cada construto e o grau de confiabilidade de cada construto para a estimação das relações causais (HAIR *et al.* 1999).

O primeiro é o de medida de indicadores exógenos, que expressa a combinação linear entre os indicadores (variáveis observáveis) e os construtos exógenos (nesses não existem flechas apontando a esses construtos). Sua forma matricial abreviada é definida por:

$$X = \Lambda^x\xi + \Theta_\delta \quad (29)$$

em que:

X (p x 1): é o vetor das indicadores do construto exógeno;

Λ^x (p x m): é o vetor de saturações, que expressa a dependência entre indicadores e construtos exógenos;

ξ (m x 1): é o vetor de variáveis latentes exógenas; e

Θ_δ (p x 1): é o vetor de resíduos, decorrentes de perturbações aleatórias e/ou sistemáticas.

Sub-índice p expressa o número de indicadores dos construtos exógenos.

O segundo é o modelo de medida de indicadores endógenos, que expressa a combinação linear entre os indicadores e os construtos endógenos (nesses existem flechas, apontando a esses construtos). Sua forma matricial abreviada é dada pelo modelo abaixo:

$$Y = \Lambda^y \eta + \Theta_\varepsilon \quad (30)$$

em que:

Y ($q \times 1$): é o vetor dos indicadores do construto endógenas;

Λ^y ($q \times n$): é o vetor de saturações, que expressa a dependência entre indicador e construtos exógenos;

η ($n \times 1$): é o vetor de variáveis latentes endógenas; e

Θ_ε ($q \times 1$): é o vetor de resíduos, decorrentes de perturbações aleatórias e/ou sistemáticas.

O sub-índice q expressa o número de indicadores dos construtos endógenos.

6.3.3 Modelo de medida de indicadores exógenos e endógenos

O modelo de medida é um conjunto de equações lineares, definido pelo número de indicadores apresentados no diagrama de caminho. A sua formulação decorre da transição da análise fatorial exploratória, onde não se tem o controle pleno sobre quais indicadores descrevem cada fator comum, para um modo confirmatório em que se especifica que indicadores definem cada fator comum.

Assim, o modelo de medida especifica um conjunto de indicadores, destinado para medir determinado fator comum. Caso o instrumento de avaliação seja projetado para medir

vários construtos, têm-se conjuntos específicos de indicadores intercorrelacionados, para cada construto. A suposição de que cada conjunto de indicadores intercorrelacionados seja independente um do outro é fundamental para o estabelecimento da validade plena do construto.

O diagrama de caminho, exposto no diagrama 2, apresenta somente construtos exógenos, possibilitando definir somente o modelo de medida de indicadores exógenos, ficando assim estabelecido:

$$X_{40} = \lambda_{40,1}\xi_1 + \delta_{40}$$

$$X_{41} = \lambda_{41,1}\xi_1 + \delta_{41}$$

$$X_{42} = \lambda_{42,1}\xi_2 + \delta_{42}$$

$$X_{49} = \lambda_{49,1}\xi_1 + \delta_{49}$$

$$X_{50} = \lambda_{50,1}\xi_1 + \delta_{50}$$

$$X_{46} = \lambda_{46,2}\xi_2 + \delta_{46}$$

$$X_{47} = \lambda_{47,2}\xi_2 + \delta_{47}$$

$$X_{48} = \lambda_{48,2}\xi_2 + \delta_{48}$$

$$X_{37} = \lambda_{37,3}\xi_3 + \delta_{37}$$

$$X_{38} = \lambda_{38,3}\xi_3 + \delta_{38}$$

$$X_{39} = \lambda_{39,3}\xi_3 + \delta_{39}$$

$$X_{45} = \lambda_{45,3}\xi_3 + \delta_{45}$$

O modelo de medida de indicadores exógenos, na forma matricial explícita, adquire a seguinte forma:

$$\begin{bmatrix} X_{40} \\ X_{41} \\ X_{42} \\ X_{49} \\ X_{50} \\ X_{46} \\ X_{47} \\ X_{48} \\ X_{37} \\ X_{38} \\ X_{39} \\ X_{45} \end{bmatrix} = \begin{bmatrix} \lambda_{40,1} & 0 & 0 \\ \lambda_{41,1} & 0 & 0 \\ \lambda_{42,1} & 0 & 0 \\ \lambda_{49,1} & 0 & 0 \\ \lambda_{50,1} & 0 & 0 \\ 0 & \lambda_{46,2} & 0 \\ 0 & \lambda_{47,2} & 0 \\ 0 & \lambda_{48,2} & 0 \\ 0 & 0 & \lambda_{37,3} \\ 0 & 0 & \lambda_{38,3} \\ 0 & 0 & \lambda_{39,3} \\ 0 & 0 & \lambda_{45,3} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} + \begin{bmatrix} \delta_{40} \\ \delta_{41} \\ \delta_{42} \\ \delta_{49} \\ \delta_{50} \\ \delta_{46} \\ \delta_{47} \\ \delta_{48} \\ \delta_{37} \\ \delta_{38} \\ \delta_{39} \\ \delta_{45} \end{bmatrix}$$

$\mathbf{X} \qquad \qquad \mathbf{\Lambda} \qquad \qquad \mathbf{\xi} \qquad \qquad \mathbf{\Theta}_\delta$

Em sua forma matricial explícita, adquire a seguinte formulação:

$$X = \Lambda^x \xi + \delta \quad (31)$$

em que

$X_{(p \times 1)}$ é o vetor de indicadores dos construtos exógenos;

$\Lambda^x_{(p \times m)}$ é a matriz de saturação que expressa o relacionamento entre indicadores e construtos exógenos;

$\xi_{(m \times 1)}$ corresponde ao vetor de variáveis latentes exógenos; e

$\delta_{(p \times 1)}$ é o vetor de resíduos ou perdas decorrentes de perturbações.

Dado que, em geral, o indicador nunca será um reflexo exato da informação dos construtos teóricos, os termos residuais, que aparecem no modelo (31) incluem os dois tipos de efeitos, devido à especificidade dos indicadores e do erro aleatório. Ambos efeitos assumem não correlacionados entre si e não correlacionados com os construtos exógenos ou endógenos.

Tendo estabelecido o modelo de medida de indicadores exógenos, segue-se para a elaboração da matriz de covariância ou correlações existentes entre os construtos e entre os indicadores. A correlação entre os construtos exógenos é representada pela matriz Φ e seus elementos por $\Phi_{m \ n}$. A matriz de correlação entre os construtos exógenos, apresentado no diagrama 2, pode ser estabelecido como segue:

$$\Phi = \begin{pmatrix} & \xi_1 & \xi_2 & \xi_3 \\ \xi_1 & - & \Phi_{12} & \Phi_{13} \\ \xi_2 & \Phi_{21} & - & \Phi_{23} \\ \xi_3 & \Phi_{31} & \Phi_{32} & - \end{pmatrix}$$

Em geral, as variáveis latentes exógenas, provenientes dos processos de validação de construto no âmbito educacional, tal como evidencia os artigos de Marsh (1987), Barton *et al.* (1994), Ramos (1997), Val (1994), estão correlacionados, representando uma influência compartilhada dos indicadores exógenos destinados a medir determinado construto com um outro construto. Quanto menos relacionados manifestarem-se as variáveis latentes exógenas, mais bem definidos serão os construtos. A existência de correlação entre os indicadores do modelo de medida de indicadores exógenos e/ou endógenos, pode ser analisada através da matriz de correlação de seus erros de medida.

O erro de medida, parte da variabilidade do indicador que não é explicado pela variabilidade da variável latente exógena, é decorrente das componentes específicas e aleatórias. A componente específica é inerente ao indicador e mantém-se estável através de sucessivas medidas. A sua existência é decorrente da influência de outras variáveis que não foram controladas durante o processo avaliativo. Por exemplo, um aluno que tenha um interesse antecipado pela disciplina de determinado professor, pode considerá-lo excelente num processo avaliativo, mesmo que os demais alunos, que estão tendo a primeira oportunidade de investigarem aquela disciplina manifestarem que falta clareza por parte do professor.

Assim, a parte específica do erro é explicada por outra variável, que nesse caso é o interesse antecipado pela disciplina. Desse modo, pode-se afirmar que o indicador está medindo aspectos de outra natureza, não sendo totalmente válido como agente do construto teórico. O erro aleatório é decorrente da falta de estabilidade quando se realizam medições repetidas com os mesmos indivíduos, sobre a mesma disciplina.

Na fase de especificação do modelo de medida, além de identificar as correlações entre os construtos exógenos e entre os erros de medida para os indicadores exógenos (θ -

delta), o pesquisador deve estar atento quanto ao número de indicadores para compor o construto, bem como, com a sua confiabilidade.

Segundo Hair *et al.* (1999), a quantidade de indicador a ser utilizado, está diretamente relacionada com a dimensão dos fundamentos teóricos que justifiquem a sua existência. Não existe um limite superior para o número de indicadores exógenos ou endógenos. Para efeitos práticos, as maiorias dos construtos deveriam ser representadas por um número entre cinco e sete indicadores saturados fortemente com o construto.

Uma vez estabelecido o modelo de medida, a análise da consistência interna do construto, mediante o coeficiente *alfa de Cronbach*, definirá o grau de confiabilidade com que os indicadores do construto exógeno e/ou endógeno indicam o construto teórico (HAIR *et al.* 1999).

6.4 Identificação do modelo de medida

Nessa etapa, a matriz de entrada-correlação ou variância-covariância das variáveis observadas deve ser investigada quanto à sua capacidade para estimações únicas dos parâmetros não conhecidos do modelo, ou seja, investigar a sua capacidade de identificação.

A título de exemplificação, será analisado se o modelo de medida do diagrama 2 é identificado. Para fornecer resposta a esse exemplo, é necessário comparar os valores conhecidos e os valores não conhecidos. Esse modelo apresenta 15 valores não conhecidos, que são as cargas fatoriais, isto é: $\lambda_{40,1}$, $\lambda_{41,1}$, $\lambda_{42,1}$, $\lambda_{49,1}$, $\lambda_{50,1}$, $\lambda_{46,2}$, $\lambda_{47,2}$, $\lambda_{48,2}$, $\lambda_{37,3}$, $\lambda_{38,3}$, $\lambda_{39,3}$, $\lambda_{45,1}$, e as correções entre os fatores latentes ϕ_{12} , ϕ_{23} e ϕ_{13} , por outro lado, existem 78 valores conhecidos que são os elementos distintos da matriz de variâncias-covariâncias dos indicadores X_{40} , X_{41} , X_{42} , X_{49} , X_{50} , X_{46} , X_{47} , X_{48} , X_{37} , X_{38} , X_{39} e X_{45} . Como existem menos parâmetros desconhecidos a serem estimados, ou seja, quinze, enquanto existem 78

estatísticas conhecidas, diz-se que o modelo de medida do diagrama 2 é sobre-identificado, ou seja, apresenta mais informações na matriz de dados de entrada que o número de parâmetros a serem estimados, significando que tem um número positivo de graus de liberdade. Dado a sua capacidade de generalização, o modelo sobre-identificado é o mais adequado para alcançar soluções isentas de restrições.

Segundo Jöreskog (1978), o número de graus de liberdade de um modelo de medida é calculado por:

$$gl = \frac{1}{2} p(p+1) - t \quad (32)$$

em que :

p corresponde ao número de indicadores endógenos (n) e exógenos (m) do modelo de medida, t corresponde ao número de coeficientes a serem estimados.

Considerando que o diagrama de caminho, apresentado na diagrama 2, não apresenta construtos endógenos, o seu grau de liberdade fica assim estabelecido:

$$gl = \frac{1}{2}[(0 + 12) (0 + 12 + 1)] - 15 = \frac{1}{2} (12*13) - 15 = 78 - 15 = 63$$

Como o modelo de medida, apresentado na diagrama 2, apresenta um número positivo de graus de liberdade, é possível estimar um modelo de medida com indicadores de ajuste aceitável com o maior grau de liberdade possível.

6.5 Estimação do modelo de medida

Antes de iniciar o processo de estimação, deve-se certificar da normalidade multivariada dos dados. A falta de normalidade multivariada é particularmente problemática porque infla substancialmente o resultado da estatística χ^2 (Qui-Quadrado), usada para o teste

estatístico de aderência dos dados. Como esse teste está baseado em uma distribuição normal, a falta de normalidade conduzirá a um aumento na probabilidade de rejeição da hipótese “de aderência dos dados”, sendo essa hipótese verdadeira. A falta de normalidade também cria um viés em valores críticos, da estatística t , usada no estudo da significância dos coeficientes.

Com relação ao tamanho da amostra, Hair *et al.* (1999) dizem que não existe nenhum critério quanto ao tamanho da amostra necessária para se realizar uma modelagem adequada, mas ressaltam a sua importância na estimação do erro amostral e de que a má especificação do modelo, aliado ao tamanho do modelo, falta de normalidade multivariada e os procedimentos de estimação são fatores que devem ser levados em consideração na determinação do tamanho da amostra necessária para a estimação dos coeficientes e indicadores do grau de ajuste de um modelo de medida. Recomendam que, para cada coeficiente a ser estimado, seja apropriado utilizar 10 avaliadores. Por exemplo, o instrumento de avaliação de disciplinas da UNISUL é composto por 29 indicadores, que se supõe estarem bem saturados em suas respectivas dimensões. Vinte e nove saturações e 28 correlações entre dimensões, serão estimadas e testadas a partir de uma amostra de 570 unidades de análises. Por outro lado, a falta de normalidade multivariada dos dados eleva de 10 para 15 o número de avaliadores por parâmetro a ser estimado (HAIR *et al.* 1999).

6.5.1 Eleição do tipo de matriz

Os modelos de equações estruturais podem usar como dados de entrada à matriz de variâncias-covariâncias ou à matriz de correlações das indicadoras (JÖRESKOG, 1978).

O uso da matriz de variância-covariância requer a construção de uma matriz quadrada, onde os elementos da diagonal principal correspondam às variâncias das indicadoras medidas, enquanto as demais posições evidenciam a covariância entre os pares de

indicadores. Covariância positiva indica que duas indicadores estão correlacionadas positivamente, enquanto covariância negativa indica duas indicadores estão correlacionadas negativamente. Por outro lado, o uso da matriz de correlação exige a elaboração de uma matriz quadrada, onde os elementos da diagonal principal são todos iguais a “um” (variância de uma indicadora padronizada), enquanto as demais posições evidenciam os correlacionamentos entre duas indicadores, os quais podem variar entre -1 e 1 . Quanto mais próximo de -1 ou $+1$ maior a intensidade destes correlacionamentos.

Tradicionalmente, o modelo de equações estrutural tem sido formulado para usar a matriz de variâncias-covariâncias. Não obstante, na área educacional, os trabalhos de Ramos (1997), Mateo e Fernández (1992), Val (1994) empregaram a matriz de correlações como dados de entrada com vistas à validação de construto.

Segundo Garcia e Martínez (2000), o uso da matriz de variâncias-covariâncias apresenta a vantagem de proporcionar comparações válidas entre diferentes populações e amostras, dado que os coeficientes estimados conservam a unidade de medidas dos indicadores. Por outro lado, os coeficientes devem ser interpretados em termos das unidades de medida para os construtos, dificultando, assim a sua interpretação.

O uso da matriz de correlação é mais atraente, pois as correlações oscilam entre -1 e 1 , possibilitando comparações diretas dos coeficientes dentro de um modelo. O seu uso é adequado quando o objetivo da modelagem constitui no estabelecimento do padrão das relações entre construtos teóricos (GARCIA; MARTÍNEZ, 2000).

Em síntese, Garcia e Martínez (2000), recomendam iniciar o processo de estimação do modelo de medida utilizando-se a matriz de variância-covariância para obter, posteriormente as soluções padronizadas para facilitar a interpretação dos resultados.

Quanto ao procedimento de estimação, o método de estimação de máxima verossimilhança é o mais comumente utilizado durante o processo de validação do construto.

6.5.2 Estimação dos parâmetros do modelo

Uma vez especificado o modelo de medida, escolhido o tipo de matriz de entrada dos dados, avaliada a identificação do mesmo, escolhe-se um procedimento para obter as estimativas dos parâmetros do modelo de medida. A estimação dos parâmetros do modelo de medida realiza-se de forma interativa através da minimização de uma função de resíduos entre a matriz de variância-covariância observada e a matriz de variância-covariância estimada (JÖRESKOG, 1978). Existem diversas técnicas de estimação, quais sejam: Método de Estimação por Máxima Verossimilhança (ML), Método de Estimação por Mínimos Quadrados Generalizados (GLS), Método de Estimação por Mínimos Quadrados Não Ponderados (ULS), Métodos de Estimação por Mínimos Quadrados Ponderados (WLS), etc.

Assumindo que a distribuição amostral dos indicadores do modelo de medida seja normal multivariada as seguintes funções de ajuste foram propostas por Jöreskog (1978):

- 1) Método dos Mínimos Quadrados Não Ponderados (ULS)

$$U(\theta) = \frac{1}{2} \text{traço} (S - \Sigma)^2 \quad (33)$$

- 2) Método dos Mínimos Quadrados Generalizados (GLS)

$$G(\theta) = \frac{1}{2} \text{traço} (I - S^{-1}\Sigma)^2 \quad (34)$$

- 3) Método de Máxima Verossimilhança (ML)

$$M(\theta) = \frac{1}{2} \text{traço} [(S - \Sigma).\Sigma^{-1}]^2 \quad (35)$$

em que:

S : é a matriz de variância-covariância amostral;

Σ : é a matriz de variância-covariância estimada pelo modelo de medida; e

I: é a matriz identidade.

As três funções **U**, **G** e **M** podem ser minimizadas, em relação à θ , basicamente pelo mesmo algoritmo. A notação $F=F_{\theta}(\mathbf{S}, \Sigma)$ poderá ser usada para qualquer uma das três funções, pois é função também dos parâmetros.

6.5.2.1 Métodos de estimação

O método de estimação por máxima verossimilhança (**ML**) é a função de ajuste mais usual em modelos de medida com vistas à validação de construto. Método de estimação usado por Ramos (1997), Val (1994), Mateo e Fernández (1992) e Barton *et al.* (1994), para confirmar as características multidimensionais da qualidade dos docentes universitários.

Segundo Bentler e Bonett (1980), a expressão de máxima verossimilhança que minimiza θ é definido por:

$$M(\theta) = \log |\Sigma| + \text{traço}(\mathbf{S}\Sigma^{-1}) - \log |\mathbf{S}| - \mathbf{K} \quad (36)$$

e que **K** corresponde à quantidade de indicadores exógenos e endógenos do modelo.

Os estimadores obtidos por máxima verossimilhança são não viesados e eficientes porém, são sensíveis ao tamanho da amostra, cujo tamanho ideal deve estar entre 100 e 200 unidades de análise (Marsh *et al.* 1988). O método de máxima verossimilhança não é recomendado para amostras superiores a 400 avaliadores, uma vez que, mesmo não existindo diferenças estatisticamente significativas entre os distintos elementos da matriz **S** e Σ , pode resultar, em uma função, perda significativa; também o contrário pode ocorrer (HAIR *et al.* 1999).

A técnica de minimização dos mínimos quadrados generalizados (**GLS**), também, pode ser utilizada para estimar os parâmetros de um modelo de medida. Comparada com a

técnica **ML**, a técnica **GLS** não apresenta sensibilidade elevada quando se utilizam amostras maiores no processo de estimação. Estimam-se estimadores não enviesados e eficientes dos parâmetros do modelo. A sua utilização com amostras pequenas podem resultar estimadores com pequeno viés. Também requer que os indicadores do modelo apresentem uma distribuição multinormal e que a matriz de entrada seja definida como positiva (JÖRESKOG, 1978).

Rejeitando-se a hipótese de normalidade multivariada dos indicadores do modelo de medida, a técnica de mínimos quadrados não ponderados (**ULS**) pode ser utilizada para estimar os parâmetros do modelo de medida. Segundo Garcia e Martínez (2000), os estimadores obtidos são pouco eficientes e sua função varia em função do tipo de matriz de entrada dos dados.

6.5.2.2 Processos de estimação

Segundo Hair *et al.* (1999), diversos processos de estimação podem ser utilizados, quais sejam: o de estimação direta, *bootstrapping*, simulação e análises de JackKnife. O processo de estimação direta faz uso de uma única amostra para estimar um modelo através da técnica de estimação escolhida, o que o torna o mais comumente usado. Do processo de minimização, resultam as estimativas dos parâmetros do modelo de medida e seus respectivos erros padrões. Por *bootstrapping* são tomadas várias amostras aleatórias dos dados originais e, para cada uma delas, um modelo de medida é estimado. Da análise conjunta de todas as estimativas para um dado parâmetro populacional, obtém-se a média e o desvio padrão que serão utilizados para prever o verdadeiro parâmetro populacional. A vantagem desse processo de estimação reside no fato de não precisar de uma distribuição de probabilidade teórica para fixar a margem de erro para as estimativas, pois a média de um dado coeficiente do modelo é

o próprio parâmetro populacional, o mesmo acontece com o desvio padrão dessas estimativas. A desvantagem é que, para estimar o verdadeiro modelo de medida, necessita-se gerar milhares de sub-amostras aleatórias de uma mesma amostra original. O processo de estimação por simulação envolve amostras múltiplas para a estimativa dos parâmetros do modelo de medida, tal como ocorre com o processo de estimação “*bootstrapping*”, distinguindo-se pela possibilidade de poder criar matrizes de entrada de dados, baseando-se em parâmetros específicos que reflitam a variação da distribuição de entrada dos dados.

O processo de estimação de JackKnife, também, exige para a sua operacionalização a criação de amostras repetidas a partir da amostra inicial. Por esse método, toda vez que se planeja uma nova amostra, omite-se uma observação diferente. Portanto, cada nova amostra é de tamanho de $n-1$ com uma observação diferente omitida em cada amostra. Permite a estimação empírica para os níveis de confiança dos parâmetros, além de possibilitar a identificação de observações influentes.

6.6 Avaliação e interpretação do modelo de medida

Os instrumentos de avaliação da docência pelos discentes usados pelas instituições de ensino superior devem, *a priori*, ser multidimensional. Como foi visto no capítulo 2, existe entre os estudiosos um relativo consenso sobre os indicadores unidimensionais que definem os bons professores e o bom ensino. A utilização dos modelos de análise fatorial confirmatória para confirmar a validade desses indicadores unidimensionais como agentes de seus respectivos construtos é realizada ajustando os dados de entrada ou seja, a matriz de correlação amostral a um modelo de medida e, a seguir, através do diagnóstico da adequação do ajuste alcançado, dotá-lo de certo grau de validade de construto.

O diagnóstico de bondade do ajuste do modelo de medida é crucial para o estabelecimento da validade da estrutura multidimensional do instrumento, pois através desses indicadores que é possível dotar o instrumento de certo grau de validade de construto. Nessa etapa de diagnóstico, é possível distinguir entre os modelos de medida que se encaixem bem nos dados, daqueles que não se encaixem bem nos dados, com o menor erro possível.

Como em todo processo de estimação, estar-se efetuando inferência sobre os verdadeiros parâmetros populacionais com um dado nível de confiança, nunca será possível afirmar que um modelo de medida seja totalmente correto e tampouco demonstrar que seja totalmente incorreto. Como afirma Ramos (1986), a validez de um instrumento é uma questão de grau.

Um modelo de medida correto incorpora as restrições e pressupostos que se cumprem na população, sem a omissão de parâmetros importantes. O modelo adequado resulta do compromisso entre a sobriedade do pesquisador e dos índices de ajuste do modelo (HAIR *et al.* 1999).

Antes de proceder ao diagnóstico da adequação do ajuste, Hair *et al.* (1999) recomendam o exame detalhado da solução obtida, objetivando a detectar problemas sérios, tais como a presença de estimações infratoras, variâncias e erro padrão negativo ou a falta de convergência do algoritmo de estimação. Resolvidos esses problemas, avalia-se globalmente o modelo de medida construído e, por último, recomenda-se diagnóstico minucioso, de cada uma das estimativas, com o objetivo de detectar partes do modelo cuja especificação não seja correta.

A presença de estimações infratoras, ou seja, aquelas que excedam seus limites teóricos, podem ser constatadas pela existência de variância residual negativa ou variância do erro significativo para qualquer indicador, coeficientes padronizados que estão muito próximo da unidade, ou erros padrões elevados dos coeficientes estimados. A correção do modelo de

medida começa pela reavaliação da identificação do modelo. Se necessário, fixa-se às variâncias dos erros infratoras um valor muito pequeno, e, no caso de coeficientes padronizados muito próximo da unidade, optar pela eliminação do indicador ou considerar que 100% da sua variabilidade seja explicada pelo construto ou dimensão latente (HAIR *et al.* 1999).

Estabelecido que o modelo está isento de estimativas infratoras, passa-se à avaliação da qualidade do modelo estimado. Os modelos de análise fatorial confirmatória não possuem um único teste estatístico para avaliar a força do modelo construído. Foram desenvolvidas várias medidas de adequação do ajuste que, quando forem usadas em conjunto, oferecem o suporte empírico para a validade de construto. São eles: as medidas de ajuste global do modelo, as medidas incrementais de ajuste e as medidas de ajuste de parcimônia (HAIR *et al.* 1999).

6.6.1 Medidas de ajuste absoluto

As medidas de ajuste absoluto determinam o grau em que o modelo de medida construído é capaz de prever com o menor erro possível a matriz de variância-covariância ou a matriz de correlação utilizada na modelagem. Dentre as principais medidas de ajuste absoluto encontrado na literatura, há o teste de χ^2 (Qui-Quadrado) citado por Marsh, Balla e McDonald (1988), a medida de parâmetro de não centralidade (PNC) proposto por McDonald e Marsh (1990), o índice de adequação do ajuste (GFI) proposto por Jöreskog e Sörbom (1984, *apud* MULAİK *et al.* 1989) e o índice quadrático médio (RMSR) proposto Jöreskog e Sorbom (1984, *apud* MARSH H.W. *et al.* 1988), que serão utilizados para medir o grau com que o modelo global construído representa a matriz de entrada dos dados.

6.6.1.1 Teste do χ^2

É a medida de adequação do ajuste mais utilizada, embora não seja capaz de funcionar adequadamente para qualquer tamanho de amostra. Sobre a hipótese de que os indicadores de ajuste de medida sejam multinormal, as seguintes hipóteses são formuladas para comparar a matriz de entrada e de saída dos dados:

Hipótese nula (H_0): $S = \Sigma$

Hipótese alternativa (H_a): $S \neq \Sigma$

em que S é a matriz de variância-covariância ou correlação observada e Σ é a matriz de variância-covariância ou correlação estimada. Um valor elevado do χ^2 , para um dado número de graus de liberdade, significa que as matrizes observadas e estimadas diferem consideravelmente. Assim, os valores de χ^2 pequenos, que resultam em p-valor $>$ que 0,05, indicam que as matrizes S e Σ são estatisticamente idênticas. Uma restrição, quanto ao uso da prova χ^2 , diz respeito à sua aplicabilidade para modelos construídos a partir de grandes amostras. O seu uso é recomendado quando o tamanho amostral está entre 100 e 200 medidas (MARSH; BALLA; MCDONALD, 1988).

6.6.1.2 Parâmetros de não centralidade

O parâmetro de não centralidade, proposto por McDonald e Marsh (1990), corresponde ao valor do χ^2 corrigido pelo número de graus de liberdade. O valor do parâmetro de não centralidade da escala é definido por:

$$\text{PNC} = \chi^2 - \text{gl} \quad (37)$$

O número de graus de liberdade (gl) para um modelo proposto é dado por:

$gl = \frac{1}{2} p(p+1) - t$, em que p corresponde ao número de indicadores exógenos (n) e exógenos (m) do modelo de medida, e t corresponde ao número de parâmetros estimados do modelo de medida.

Como o PNC ainda é influenciado pelo tamanho da amostra, McDonald e Marsh (1990) propõem dividir o PNC pelo tamanho da amostra. Essa medida denominada de “Parâmetro de Não Centralidade da Escala (PNCE)” é também especialmente útil para comparar modelos alternativos e, quanto menor, for o seu valor, melhor é a qualidade do ajuste. Os indicadores da qualidade do ajuste “PNC” e “PNCE” são recomendados para comparar modelos de medidas alternativos e se escolhe o modelo que apresentar valor mais próximo de zero.

6.6.1.3 Índice de adequação do ajuste (GFI)

O índice de adequação do ajuste “GFI”, proposto por Jöreskog e Sörbom (1984, *apud* MULAİK *et al.* 1989), é um índice da variabilidade explicada pelo modelo construído pelo método de máxima verossimilhança. O índice GFI é obtido por:

$$GFI = 1 - \left\{ \frac{\text{traço}(\Sigma^{-1} S - I)^2}{\text{traço}(\Sigma^{-1} S)^2} \right\} \quad (38)$$

em que :

Σ : é a matriz de variância-covariância predida pelo modelo;

S : é a matriz de variância-covariância de entrada;

Traço: é a soma dos elementos da diagonal principal; e

I : é a matriz identidade.

O índice GFI é um indicador do poder explicativo do modelo. Pode se manifestar super valorado, uma vez que para o seu cálculo não se levam em consideração o tamanho da amostra e o número de graus de liberdade. O GFI é um número entre zero (0) e um (1). Valores superiores a 0,90 indicam um ajuste aceitável.

6.6.1.4 Resíduo quadrático médio (RMSR)

Jöreskog e Sorbom (1984, *apud* MARSH, *et al.* 1988), definem o RMSR como a raiz quadrada da soma dos resíduos quadrados, ou seja, uma média dos resíduos observados entre as matrizes estimada e observada. Quando Σ e S são baseadas na matriz de correlação, então o RMSR está em termos de uma correlação residual média. O RMSE pode ser obtido pela seguinte expressão:

$$\text{RMSR} = \{2\Sigma\Sigma(S_{ij} - \sigma_{ij})^2 / [p(p+1)]\}^{1/2} \quad (39)$$

em que S_{ij} são os elementos da matriz de correlação observada, σ_{ij} são os elementos da matriz de correlação estimada e p corresponde ao número de indicadores exógenos e endógenos do modelo de medida. Quando a qualidade do ajuste for bem estabelecida, o valor de RMSR estará próximo de zero. Esse indicador torna-se muito importante, quando existir a necessidade de comparar modelos alternativos.

6.6.2 Medidas incrementais de ajuste

Medidas incrementais de ajuste são indicadores da qualidade do ajuste, que possibilitam comparar o modelo estimado com um modelo teórico nulo, ou seja, visa determinar se todos os indicadores estão associados a um único fator latente. Dentre as

principais medidas incrementais de ajuste citada na literatura, há o índice de adequação corrigido (AGFI) proposto por Jöreskog 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).

6.6.2.1 Índice de adequação corrigido (AGFI)

Jöreskog e Sorbom (1981, *apud* MARSH, *et al.* 1988) apresentam esse índice com sendo uma extensão do índice GFI, ajustado pela razão entre os números de graus de liberdade do modelo estimado pelo método da máxima verossimilhança e dos graus de liberdade do modelo nulo. O índice de adequação corrigido (AGFI) é definido por:

$$AGFI = 1 - [p(p+1) / 2gl] \cdot (1 - GFI) \quad (40)$$

em que p corresponde ao número de indicadores exógenos e endógenos, gl é o número de graus de liberdade do modelo nulo, ou seja, $gl = \frac{1}{2} p(p+1) - t$, em que t corresponde ao número de parâmetros do modelo nulo. O índice AGFI é recomendado por ser independente do tamanho da amostra e indica modelo adequado quando superar 0,9 (HAIR *et al.* 1999).

6.6.2.2 Índice de Tucker-Lewis (TLI)

O índice incremental de ajuste de Tucker e Lewis (1973) resulta da comparação do número de graus de liberdade do modelo estimado com o número de graus de liberdade do modelo nulo. Ao considerar os graus de liberdade dos modelos, o problema de modelos sobreajustados como consequência do número de coeficientes estimados fica atenuado. O índice de Tucker e Lewis (TLI) é definido por:

$$TLI = \{[(\chi^2_{(nulo)}/gl_{(nulo)}) - (\chi^2_{(proposto)}/gl_{(proposto)})] / (\chi^2_{(nulo)}/gl_{(nulo)} - 1)\} \quad (41)$$

O TLI é pouco influenciado pelo tamanho da amostra. Modelos ajustados com valores de TLI superiores a 0,90 são considerados como bons ajustes (HAIR *et al.* 1999).

6.6.2.3 Índice de ajuste normal (BBI)

O índice de ajuste normal, proposto por Bentler e Bonnett (1980) compara a redução proporcional em função do ajuste quando se confronta o modelo nulo com o modelo proposto. O BBI é definido por:

$$BBI = (\chi^2_{(nulo)} - \chi^2_{(proposto)}) / \chi^2_{(nulo)} \quad (42)$$

O BBI varia entre zero (0) e um (1), sendo recomendado valores superiores a 0,90 (HAIR *et al.* 1999).

6.6.3 Medidas de ajuste de parcimônia

Um modelo é parcimonioso quando não contém coeficientes desnecessários, ou seja, é um modelo simples, mas com grande capacidade explicativa. As medidas de ajuste de parcimônia relacionam a bondade do modelo com o número de coeficientes necessários para alcançar esse nível de ajuste. As principais medidas de ajuste de parcimônia são: 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), qui-quadrado normalizado, proposto por Jöreskog (1969, *apud* GARCIA; MARTINEZ, 2000), e critério de informação de Akaike (1987).

6.6.3.1 Índice de ajuste normalizado de parcimônia (PNFI)

Discutido por Mulaik. *et al.* (1989), o PNFI corrige o índice de Bentler e Bonnett (1980), levando em consideração os graus de liberdade do modelo nulo e proposto. Sua principal aplicação consiste em comparar modelos com graus de liberdade distintos. A parcimônia fica estabelecida quando se obtém melhores níveis de ajuste por grau de liberdade utilizado. O PNFI é obtido por:

$$\text{PNFI} = \{ \text{gl}_{(\text{proposto})} / \text{gl}_{(\text{nulo})} \}. \text{BBI} \quad (43)$$

Segundo Hair *et al.*, (1999), quanto maior o valor do PNFI, maior a parcimônia e que diferenças de 0,06 a 0,09, nos valores do índice de parcimônia “PNFI”, são indicativos de diferenças substanciais entre os modelos.

6.6.3.2 Índice de bondade de ajuste de parcimônia (PGFI)

O índice de qualidade de ajuste de parcimônia (PGFI), proposto por Mulaik *et al.*, (1989), corrige o índice de bondade de ajuste (GFI) pela incorporação da parcimônia do modelo estimado pelo método da máxima verossimilhança. Ajusta o GFI da seguinte forma:

$$\text{PGFI} = \{ 2\text{gl}_{(\text{proposto})} / p(p+1) \} \times \text{GFI} \quad (44)$$

em que $\text{gl} = \frac{1}{2} p(p+1) - t$, p corresponde ao número de indicadores exógenos e endógenos e t corresponde ao número de coeficientes estimados do modelo de medida. O PGFI penaliza os modelos com muitos parâmetros e resultam em medidas que variam entre 0 e 1 e, quanto mais próximo de 1, maior a parcimônia do modelo.

6.6.3.3 Qui-quadrado normalizado

Essa medida proposta por Jöreskog (1969, *apud* HAIR *et al.* 1999) consiste no valor do teste de qui-quadrado dividido pelo número de graus de liberdade. Embora seja sensível ao tamanho da amostra, oferece duas formas de avaliar modelos inadequados: aqueles modelos sobreajustados, ou seja, aqueles cujos coeficientes padronizados estão próximo da unidade e aqueles que não apresentam um ajuste suficiente aos dados, ou seja, aqueles cujos coeficientes padronizados estão próximo de zero . Sua formulação é:

$$\chi^2_{(\text{normalizado})} = \chi^2 / gl \quad (45)$$

O número de graus de liberdade do modelo avaliado é definido por $gl = \frac{1}{2} p(p+1) - t$, em que p corresponde ao número de indicadores exógenos e endógenos e t corresponde ao número de coeficientes estimados do modelo de medida. O modelo é considerado adequado quando o qui-quadrado normalizado estiver entre 1 e 2.

6.6.3.4 Critério de informação de Akaike (AIC)

O critério de informação de Akaike proposto por Akaike (1987), serve para comparar modelos que apresentam diferentes números de construtos. O AIC pode ser calculado como:

$$AIC = \chi^2 + 2.p \quad (46)$$

em que p corresponde ao número de parâmetros estimados. Valores de AIC pequenos e positivos indicam alta parcimônia. É útil para comparar modelos alternativos (HAIR *et al.* 1999).

Segundo Bentler (1978, apud RAMOS, 1987) ao assegurar a validade de construto por meio de modelos causal, o grau de ajuste entre um modelo causal e os dados é o indicativo do grau de validade de construto do instrumento de avaliação. Apesar da grande quantidade de índices apresentados, não existe consenso sobre qual é o melhor índice de qualidade do ajuste. Assim, o melhor que se faz é combinar todos esses indicadores e escolher o modelo de medida que combine os melhores índices de ajustes absoluto, incremental e de parcimônia.

O quadro 4, apresenta uma síntese dos indicadores de ajustes absoluto, incremental e de parcimônia utilizados para a avaliação da adequação do ajuste do modelo, adaptado de Hair et al (1999). O grau da validade dos construtos operacionalizados, no instrumento de avaliação, é função dos níveis alcançados por esses indicadores.

Medidas de Adequação do Ajuste	Utilização	Níveis de Aceitação Recomendados
Medidas de Ajuste absoluto		
Estatística de χ^2	Comprovar a significância do teste.	$p_{(valor)} > 0,05$
Parâmetro de não centralidade	Comparar modelos alternativos.	Modelo que apresentar menor PNC.
Parâmetro de não centralizado de escala	Comparar modelos alternativos.	Modelo que mais se aproxima de zero.
Índice de Bondade de ajuste (GFI)	Poder explicativo do modelo.	Superior a 0,9.
Erro de aproximação quadrático médio. (RMSE).	Comparar modelos alternativos.	Abaixo de 0,08.
Medidas Incrementais de Ajuste		
Índice de bondade de ajuste (AGFI)	Poder explicativo do modelo corrigido.	
Índice de ajuste Tucker-Lewis (TLI)	Comparar o modelo proposto com o modelo nulo.	Superior a 0,90.
Índice de ajuste normalizado (BBI)	Comparar o modelo proposto com o modelo nulo.	Superior a 0,90.
Medidas de Ajuste de Parcimônia		
Índice de ajuste normalizado de parcimônia (PNFI)	Comparar modelos alternativos.	Superior a 0,90.
Índice de Bondade de Ajuste (PGFI)	Comparar modelos alternativos.	Diferenças entre 0,06 e 0,09 resultam importantes.
Critério de Informação de Akaike (AIC)	Comparar modelos com diferentes números de construtos	Escolhe-se o modelo que mais se aproxima de 1. Escolhe-se o modelo que mais se aproxima de zero.

Quadro 4: Resumo das medidas de adequação do ajuste

6.6.4 Ajuste do modelo de medida

Todas as medidas de ajustes absoluto, incremental e de parcimônia são empregadas durante o processo de validação de construto para estabelecer o grau de ajuste entre o modelo teórico proposto e os dados empíricos. Garcia e Martinez (2000) recomendam que, embora tenhamos bons indicadores do ajuste global do modelo de medida estimado, seja feita uma análise individual dos indicadores exógenos ou exógenos para evitar parâmetros estimados não significativos. Caso alguns dos indicadores exógenos 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. Estabelecida à significância das cargas, deve-se investigar a unidimensionalidade de cada indicador, bem como a confiabilidade composta do construto. Considera-se que, cada indicador deveria ter ao menos 50% da variância em comum com a variância da variável latente (SHARMA 1996, *apud* GARCIA; MARTINEZ, 2000).

6.6.5 Confiabilidade composta do construto.

De acordo com Hair *et al.* (1999), a confiabilidade composta do construto é uma medida que indica a consistência interna dos indicadores do construto e representa o grau em que os indicadores evidenciam o construto comum latente. Para cada construto latente, a confiabilidade composta e a variância extraída devem ser calculadas.

A confiabilidade composta de um construto (Ω), segundo Hair *et al.* (1999), calcula-se como:

$$\Omega_j = 1 - \frac{(\sum \lambda_{ij})^2}{\sum \varepsilon_j} \quad (47)$$

Em que:

λ_{ij} : é a saturação da indicadora (i), sobre o construto (j); e

ε_j : é o erro de medida, ou seja, $\varepsilon_j = 1 - \lambda_{ij}^2$.

Como se pode perceber, a confiabilidade do construto, tal como Ω_j , é uma medida da consistência interna dos indicadores e representa o grau em que eles indicam o construto comum latente. Assim, por exemplo, se a confiabilidade do construto for igual a 0,95, há o indicativo de que 95% da variabilidade do construto seja compartilhada com as variabilidades dos indicadores.

6.6.6 Variância extraída

A variância extraída é uma medida utilizada para avaliar o ajuste do modelo de medida. Ela indica quanto da variância global dos indicadores exógenos é explicado pelo construto latente. É uma medida complementar da confiabilidade do construto. A medida da variância extraída (σ_j^2) é obtida, segundo Hair *et al.* (1999), como:

$$\sigma_j^2 = 1 - \frac{\sum \lambda_{ij}^2}{\sum \varepsilon_j} \quad (48)$$

em que λ_{ij} é a saturação do indicador (i) sobre o construto (j), ε_j é o erro de medida, ou seja, $\varepsilon_j = 1 - \lambda_{ij}^2$. Em geral, a quantidade global de variância dos indicadores explicada pela variável latente deve exceder a 0,5, para se ter um construto consistentemente evidenciado.

6.7 Interpretação e modificação do modelo

Difícilmente os modelos de análise fatorial confirmatório ajustam-se aos dados em um primeiro contraste de hipóteses. A análise detalhada dos coeficientes estimados do modelo de medida de indicadores exógenos revelará se o seu tamanho e sentido estão acordados conforme a teoria, como também maneiras de melhorá-lo. Dessa forma, a modificação do modelo tem-se convertido em prática habitual para otimizar os índices de bondade de ajuste em função dos parâmetros estimados e conseguir maior parcimônia ou aumento da validade pela suspensão de indicadores que não alcançam índices de confiabilidade e validade apropriados. Assim fez Ramos (1987), eliminando do modelo de medida de indicadores exógenos, os indicadores com elevado erro de medida e neutros, visando obter um modelo consistente e parcimonioso para a escala de medida da competência docente do professor universitário. Ao suprimir do modelo de medida cinco indicadores que não apresentavam suficiente validade, os indicadores de bondade de ajuste **G.F.I.**, **A.G.F.I.**, **R.S.H.R.** alteram, respectivamente, de 0,888 para 0,928, de 0,858 para 0,900 e de 0,049 para 0,037.

O processo de modificação do modelo de medida é guiado, segundo Hair *et al.* (1999), pelas análises dos resíduos padronizados da matriz de correlação ou da matriz de variâncias-covariâncias e pelos índices de modificação.

A diferença entre os elementos da matriz de correlação ou variância-covariância de entrada (matriz observada) e de saída (matriz predita pelo modelo proposto) denomina-se resíduos ou perdas. O ideal é que esses resíduos padronizados sejam significativamente iguais a zero, o que indicaria que o modelo proposto reproduziu a matriz de entrada. Considerando um nível de significância de 1%, aqueles resíduos padronizados com valores superiores a

$\pm 2,58$, teste nas duas caudas, indicam uma diferença significativa entre os elementos da matriz observada e predita pelo modelo.

Os resíduos significativos indicam um erro de predição substancial para um par de indicadores. Segundo Hair *et al.* (1999), um resíduo padronizado indica somente se a diferença é ou não casual. Desse modo, qualquer modificação no modelo inicial proposto, deve ser respaldada pela teoria que combinada com os índices de modificação, contribuirão para um aumento no grau de validade do construto.

O índice de Modificação refere-se aos valores calculados para cada possível relação não estimada em um modelo especificado (parâmetros fixos). O seu cálculo é decorrente de saturações sobre outros fatores ou covariância entre fatores únicos em um modelo de análise fatorial confirmatória e se forem significativos, indicam invalidez. O valor do índice de modificação para uma relação concreta não estimada indica uma redução no valor do teste de χ^2 quando o modelo fora efetivamente estimado. Ressalta-se que qualquer restrição a ser feita no modelo, originará um novo modelo a ser estimado e avaliado. O cálculo da proporção de aumento no ajuste, segundo Garcia e Martinez (2000) é calculado por:

$$\Delta = (\chi^2_o - \chi^2_m) / \chi^2_o \quad (49)$$

Em que χ^2_o é o valor do teste qui-quadrado do modelo original, χ^2_m corresponde ao valor do qui-quadrado do modelo modificado. Esse índice oscilará entre zero e um, propondo-se o valor de 0,9 como o limite para paralisar o processo de reformulação do modelo, segundo Bisquerra (1989, *apud* GARCÍA; MARTÍNEZ, 2000). Recomenda-se, também, introduzir somente modificações plausíveis de maneira seqüencial, reexaminando atentamente todos os resultados antes de efetuar a seguinte e começar analisando os

coeficientes estimados significativos antes de eliminar os indicadores pouco válidos (MACCALLUM, *et al.* 1992; LUIJBEN, 1989; HAIR *et al.* 1999).

Segundo Hair *et al.* (1999), quando houver a necessidade de modificações amplas do modelo, os dados devem ser divididos em duas amostras independentes, uma delas será usada para a estimação do modelo modificado e a outra para a validação cruzada do modelo. Com esse procedimento, espera-se que os resultados e o modelo final proposto sejam generalizáveis para a população com um dado nível de confiança.

Tendo concluído que se tenha obtido razoável grau de validade de construto; ou seja, estabelecido que, cada conjunto de indicadores individuais intercorrelacionados se combina para tornar claro uma única medida composta ou indicador, com índices de validade e consistência interna razoáveis, têm-se evidências que designam o questionário de avaliação da docência pelos discentes como confiável e válido. Como o construto foi validado mediante modelos de análises fatorial confirmatória, já se tem o indicativo da validade convergente. Quanto à validade discriminante, ela estará evidenciada se os construtos teóricos se manifestarem incorrelacionados.

A validade convergente indica até que ponto duas ou mais escalas, dois ou mais procedimentos independentes que tratam de medir o mesmo conceito, convergem para a mesma solução. A validade discriminante mostra até que ponto um indicador não se correlaciona com outros construtos exógenos dos quais supõe-se que difira. Quanto menor o grau de correlação entre dois construtos exógenos, maior a evidência de validade discriminante.

Cada dimensão validada da escala pode ser correlacionada com outros indicadores previamente estabelecidos, com vistas à validade preditiva ou concorrente. O indicativo de validade preditiva será evidenciado quando as dimensões válidas da escala estiverem correlacionadas positivamente e significativamente com algum critério que globalmente

evidencia a qualidade do ensino, no futuro. Os indicativos de validade concorrente serão evidenciados empiricamente quando as dimensões válidas da escala estiverem correlacionadas positivamente e significativamente com indicadores distintos do mesmo conceito.

7 APLICAÇÃO DA METODOLOGIA AO INSTRUMENTO DE AVALIAÇÃO DA DOCÊNCIA DA UNISUL

7.1 Introdução

O presente capítulo tem como objetivo descrever os procedimentos metodológicos utilizados no processo de análise da confiabilidade e da validade do instrumento de avaliação da UNISUL. A revisão teórica sobre o estudo da dimensionalidade, da confiabilidade e da validade de instrumentos de avaliação da docência apresentada, respectivamente, nos capítulos anteriores, constitui o referencial teórico sobre a validação do instrumento de avaliação. Tal referencial possibilitará o contraste entre as características de multidimensionalidade, confiabilidade e validade do instrumento “Avaliação pelos Alunos no Desempenho na Disciplina” utilizado nessa universidade.

Este capítulo descreve sobre a entidade, o tipo de pesquisa, o objetivo da pesquisa, o método da pesquisa, indicadores medidos (tipos, coleta e tratamento), hipóteses e limitações da pesquisa.

7.2 A entidade

A Fundação Universidade do Sul de Santa Catarina – UNISUL, situada no município de Tubarão, Sul do Estado de Santa Catarina foi instituída pelo poder público municipal, tendo natureza jurídica de direito privado. É de caráter filantrópico e integra o Sistema Educacional do Estado de Santa Catarina – ACAFE.

Fundada em 1964, como IMES – Instituto Municipal de Ensino Superior, passou a ofertar um único curso de graduação: Ciências Econômicas. Em 1967, através da lei

municipal o IMES foi transformado em Fundação Educacional do Sul de Santa Catarina – FESSC. Em 1989, através de portaria ministerial foi reconhecida como universidade.

A gestão da Universidade faz-se através de um Conselho Curador e pelo Conselho Universitário. O Conselho Curador é o órgão de representação das instituições mantenedoras e de fiscalização, já o Conselho Universitário é o órgão de naturezas normativas, deliberativas, jurisdicional e consultiva de assuntos de planejamento e administração geral e de matéria de ensino, pesquisa e extensão.

Considera-se, em termos estatutários, como região de influência da Universidade do Sul de Santa Catarina – UNISUL, prioritariamente, a faixa do Estado de Santa Catarina, compreendida entre a Serra Geral e o Oceano Atlântico, a partir da capital do Estado, Florianópolis, até a divisa com o Estado do Rio Grande do Sul.

A UNISUL, hoje consolidada como Universidade Comunitária Regional, irradia suas ações através de unidades descentralizadas em sua área de influência:

- Campus I (Campus de Tubarão –sede), com extensão nos municípios de Laguna, Imbituba e Braço do Norte.

- Campus II (Campus de Araranguá), foi instalado, em 1992, em Araranguá, município distante 70 Km ao Sul de Tubarão, com extensão no Município de Içara.

- Campus III (Campus da Pedra Branca), foi instalado, em 1996 no Município de Palhoça, com extensão no Município de Florianópolis.

A tabela 11 apresenta os números de cursos e o total de alunos no ano de 2002, segundo os *campi*.

Tabela 11: Número de cursos e de alunos por nível de ensino-2002.

Nível de Ensino	Tubarão		Araranguá		Pedra Branca	
	Nº de cursos	Nº de alunos	Nº de cursos	Nº de alunos	Nº de cursos	Nº de alunos
Seqüencial	14	629	1	63	7	1305
Graduação	37	8487	15	2492	26	7345
Especialização	30	1016	6	153	18	706
Mestrado	6	512	-	-	4	360
Doutorado	7	375	-	-	1	28
Total	94	11019	22	2708	56	9744

7.3 Tipo de pesquisa e base de dados

Esta pesquisa trata da investigação das propriedades psicométricas de confiabilidade e validade do instrumento “Avaliação pelos Alunos do Desempenho na Disciplina”, utilizado atualmente pela equipe de Avaliação Institucional e Clima Organizacional da UNISUL. Espera-se que a base de dados proveniente da última pesquisa de avaliação institucional, ocorrida no 2º semestre letivo de 2002, nos três *campi*, possibilite informações relevantes e úteis para a comprovação das propriedades de confiabilidade e validade do instrumento.

A tabela 12 apresenta o total de alunos matriculados por campus no semestre letivo de 2002, bem como a quantidade de alunos que participaram da avaliação e os respectivos percentuais. Os valores apresentados entre parênteses correspondem ao número de alunos que estudam nos cursos de licenciaturas nos finais de semana que não participaram do processo de avaliação.

Tabela 12: Matriculados na UNISUL, no segundo semestre letivo de 2002.

Campus	Alunos Matriculados	Avaliação	%
Tubarão	8487 (1030)	5310	71,2
Araranguá	2492 (308)	1318	60,3
Pedra Branca	7345 (204)	4174	59,5
UNISUL	18324 (1542)	10802	64,4

Nessa avaliação, todos os alunos da UNISUL foram convidados a manifestar o seu nível de satisfação, através de quatro categorias de respostas (sim, parcialmente sim, parcialmente não e não), em 65 questões do instrumento de “Avaliação pelos Alunos do Desempenho na Disciplina”. Cada formulário foi lido pela leitora óptica, da seguinte forma:

- quando o aluno manifestar seu nível de satisfação como “completamente satisfeito” ou “sim”, a leitora atribuía o código um;
- quando o aluno manifestar seu nível de satisfação como “parcialmente satisfeito” ou “parcialmente sim”, a leitora atribuía o código dois;
- quando o aluno manifestar seu nível de satisfação como “parcialmente insatisfeito” ou “parcialmente não”, a leitora atribuía o código três; e
- quando o aluno manifestar seu nível de satisfação como “completamente insatisfeito” ou “não”, a leitora atribuía o código quatro.

A partir dessa leitura, foi aberta uma pasta para cada curso, contendo o registro da matrícula do professor, o código da disciplina e todos os indicadores avaliados pelos discentes da UNISUL. Essa é a base de dados que será utilizada na pesquisa dos indicativos de confiabilidade e validade do instrumento.

Em face do exposto, a presente pesquisa, numa primeira etapa, tem característica de pesquisa explicativa, uma vez que o principal objetivo é a elaboração de conceitos, com base nas sucessivas análises exploratórias, utilizadas para comprovar a validade discriminante dos indicadores como agentes das dimensões operacionalizadas no instrumento. As

dimensões, consistentemente estabelecidas por procedimentos correlacionais, constituem-se em uma hipótese que pode ser validada por procedimentos específicos de análise de dados confirmatórios.

Do ponto de vista da forma de abordagem do problema, esta pesquisa é qualitativa, ou seja, utiliza indicadores que se tornam quantificáveis em função da escala ser aditiva e qualitativa por requerer um referencial teórico bastante abrangente para o estabelecimento das dimensões. Quanto aos seus objetivos, a autoria entende que se trata de um estudo de caso, pois a preocupação central consiste na determinação do grau de confiabilidade e de validade dos construtos latentes que determinam ou que contribuem para a variabilidade dos distintos indicadores do instrumento. Características que são únicas do instrumento em análise.

7.4 Objeto da pesquisa

O estudo da confiabilidade e da validade das dimensões subjacentes ao instrumento “Avaliação pelos Alunos do Desempenho na Disciplina”, aspecto que constitui o núcleo central desta tese, coincide com os objetivos da equipe de avaliação institucional e clima organizacional da UNISUL. Assim, o objeto da pesquisa é o próprio instrumento de avaliação que, desde 1996, é valorado pelos alunos em sucessivas edições.

7.5 Método da pesquisa

De acordo com Coelho (2003), a investigação científica depende de um conjunto de procedimentos intelectuais e técnicos para que seus propósitos sejam alcançados. O Método científico consiste em um conjunto de processos e operações mentais adotado na

investigação, ou seja, é a linha de raciocínio adotada na execução da pesquisa (GIL, 1994; MARCONI e LAKATOS, 2002). Assim, no contexto desta investigação, o método científico consiste no conjunto de procedimentos intelectuais e técnicos, utilizado para investigar as propriedades de confiabilidade e validade das dimensões subjacentes ao instrumento em análise.

O método utilizado, na presente tese, é o hipotético-dedutivo. Segundo Rauen (2002), de acordo com esse método, faz-se necessário enunciar as hipóteses que, posteriormente, devem ser submetidas ao crivo de exaustivos testes, os quais procuram por evidências empíricas para torná-las falsas com o menor erro possível. Se a afirmação resistir a testes severos, estará corroborada, mas não confirmada, dado que a confirmação plena da hipótese é de difícil acesso, uma vez que se teria de acumular todos os dados presentes, passados e futuros. Todas as características desse método perfazem as etapas do processo de validação de construto proposto por Ramos (1986), que será utilizado para dotar o instrumento em análise de certo grau de validade de construto.

7.6 Variáveis do estudo

As variáveis do estudo são aquelas do instrumento “Avaliação pelos Alunos do Desempenho na Disciplina”, composto por cinco (5) blocos, quais sejam:

Bloco um (1): formado por 24 indicadores, numerados 01 a 24, que tratam da qualidade da UNISUL, do curso e outras.

Bloco dois (2): formado por 07 indicadores, numerados de 25 a 31, que se relacionam diretamente com a gerência do curso de graduação.

Bloco três (3): formado por 31 indicadores, numerados de 32 a 62, que procuram avaliar o desempenho do docente no âmbito da sala de aula.

Bloco cinco (5): formado por três indicadores, que tratam da autopercepção dos alunos quanto às suas responsabilidades e desempenho na disciplina.

Os indicadores do terceiro bloco foram designados para medirem, discriminadamente, oito dimensões latentes da qualidade do docente. O quadro 5, descreve as oito dimensões com a sua respectiva composição.

Dimensões	Nº do Indicador	Descrição
Exigência	40	Estimula a utilização de bibliografia atualizada.
	41	Estimula a produção científica, promovendo atividades de pesquisa na disciplina.
	42	Incentiva a leitura de livros, textos, jornais e revistas complementares às aulas.
	49	Trabalha a disciplina com nível de exigência suficiente.
	50	Conduz o ensino da disciplina com qualidade.
Avaliação	46	Redige provas ou verificações de aprendizagem de forma clara. Informa sobre os critérios adotados na avaliação.
	47	Planeja as avaliações compatíveis com os objetivos e conteúdos ministrados.
	48	
Conteúdo	37	Revela domínio de conteúdo na sua disciplina.
	38	Expressa o conteúdo das aulas com linguagem acessível aos alunos. A bagagem intelectual é estimulante para os alunos.
	39	Dá atenção às dúvidas apresentadas pelos alunos.
	45	
Método	43	Adota procedimentos facilitadores da aprendizagem.
	44	Oportuniza reflexões críticas sobre os conteúdos.
	52	Valoriza a participação dos alunos nas aulas.
Relacionamento Professor / Aluno	51	Tem boa relação com alunos.
	53	Estimula o bom relacionamento da classe.
Produtividade	32	É pontual, cumpre os horários das aulas.
	33	É assíduo, comparece às aulas.
	34	Administra o tempo em sala de aula de forma produtiva.
	35	Apresentou o plano de ensino.
	36	Está desenvolvendo o plano de ensino.
Ética	56	Demonstra interesse e compromisso com a UNISUL.
	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.
Relação Professor / Curso	54	Demonstra conhecer o projeto do curso
	55	Situa o aluno sobre a importância de sua disciplina para o curso
	57	Demonstra conhecimento da UNISUL no seu <i>campus</i> .

Quadro 5: Dimensões do bloco “Desempenho do Docente”.

7.7 Tratamento dos dados

Os indicativos de confiabilidade e de validade dependem de quão bem as avaliações desses indicadores pelos alunos estejam isentos 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 Guilford e Fruchter (1978), Hair *et al.* (1999), o uso das médias como unidade de análise atenua o de 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 (1984), 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 à confiabilidade e à validade das dimensões da qualidade do ensino medidas pelo instrumento de avaliação.

Ciente dessa conveniência, a média aritmética de todos os indicadores avaliados pelos alunos de uma mesma classe e referentes à avaliação da mesma disciplina é a unidade de análise. Assim, as propriedades de confiabilidade e validade das dimensões subjacentes ao instrumento serão estimadas a partir de uma base de dados composta por todas as disciplinas avaliadas. A base de dados é organizada em uma matriz em que cada linha da matriz corresponde uma disciplina e as colunas os indicadores avaliados.

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.

7.8 Hipóteses da pesquisa

O modelo qualitativo teórico, apresentado no fluxograma 5, que constitui a primeira hipótese a ser testada, foi elaborado a partir de consulta realizada com todos os segmentos da universidade, que indicaram as dimensões acadêmicas que melhor representam o desempenho dos docentes e quais eram os indicadores de cada uma das dimensões. Assim, a primeira hipótese a ser testada por procedimentos empíricos é:

Hipótese 1: o modelo qualitativo teórico, apresentado no diagrama 3, é composto por oito dimensões não correlacionadas consistentes e confiáveis da qualidade do ensino no âmbito da sala de aula.

O diagrama 3, corresponde ao diagrama de caminho das características da Qualidade Docente (oito fatores latentes não correlacionados).

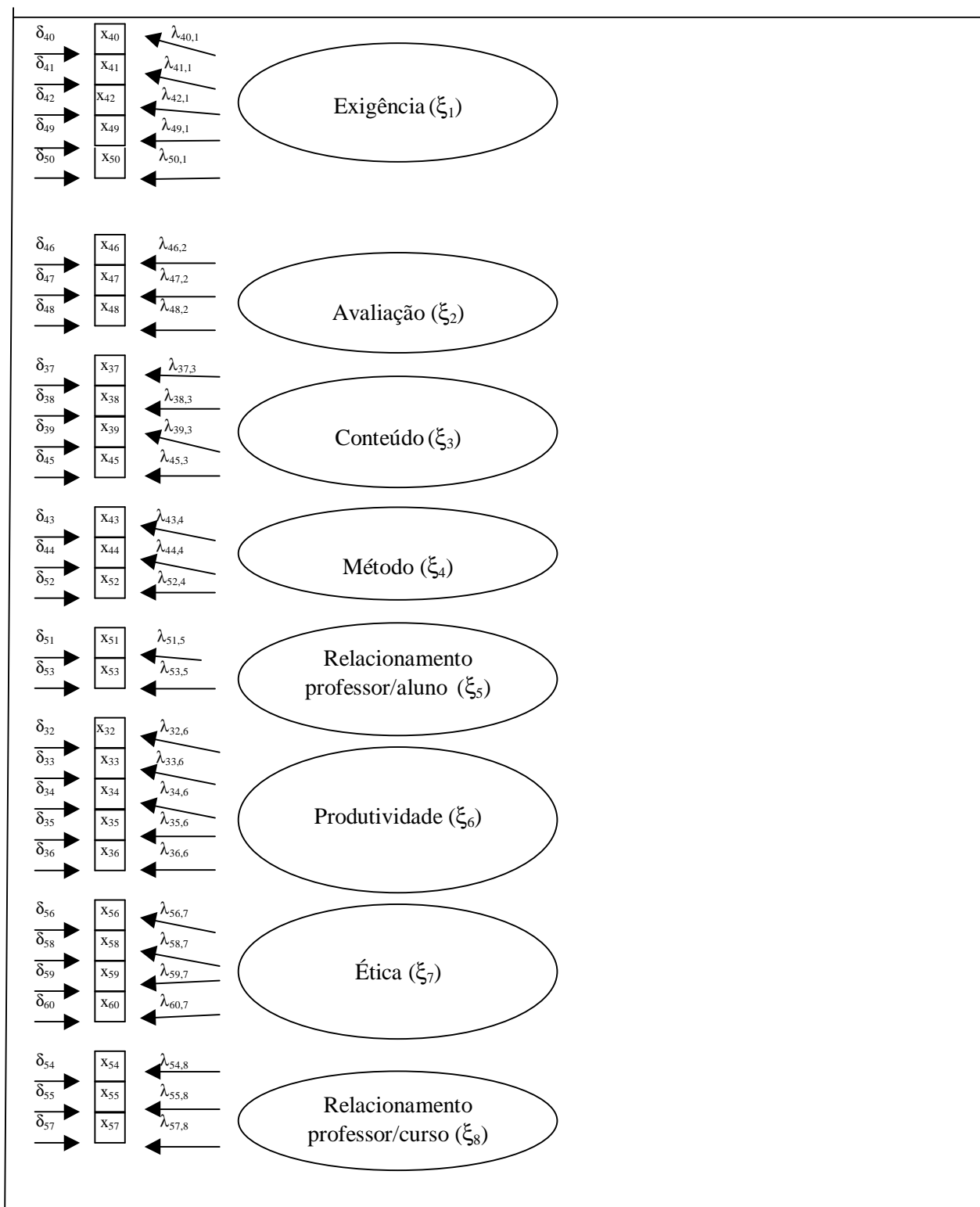


Diagrama 3: Diagrama de caminhos das características da qualidade docente (oito fatores latentes não correlacionados)

Uma dimensão é confiável e válida quando apresenta um conjunto de indicadores unidimensionais com nível de correlação bem estabelecido entre si. As sub-hipóteses 1.1, 1.2, 1.3, 1.4, 1.5 e 1.6 enunciadas abaixo, fazem-se necessárias para a comprovação dessas propriedades fundamentais do instrumento. A não rejeição dessas sub-hipóteses conduz à aceitação da hipótese um (1), e evidencia-se o suporte empírico a favor das propriedades de confiabilidade e validade das dimensões operacionalizadas no instrumento, ou seja, a validade de construto.

Sub-hipótese 1.1: as valorações estudantis atribuídas, para o total de disciplinas cursadas no 2º semestre letivo de 2002, na UNISUL, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento de avaliação, ou seja, as oito dimensões não correlacionadas do instrumento em análise.

Sub-hipótese 1.2: as dimensões não correlacionadas decorrente da análise confirmatória são consistentes, conforme os critérios psicométricos.

Sub-hipótese 1.3: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre de 2002, no *campus* de Tubarão, constituem uma base de dados para o suporte empírico das oito dimensões não correlacionadas do instrumento de medida.

Sub-hipótese 1.4: as dimensões não correlacionadas, decorrentes da análise confirmatória, para o total de disciplinas cursadas no *campus* de Araranguá, são consistentes, conforme os critérios psicométricos.

Sub-hipótese 1.5: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre de 2002, no *campus* da Pedra Branca, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento de medida, ou seja, das oito dimensões não correlacionadas do instrumento de medida.

Sub-hipótese 1.6: as dimensões não correlacionadas decorrentes da análise confirmatória, para o total de disciplinas cursadas no *campus* da Pedra Branca, são consistentes, conforme os critérios psicométricos.

A hipótese 1 estabelece que o instrumento de avaliação da docência é formado por oito (8) dimensões latentes não correlacionados. Hipótese que é difícil de ser sustentada, uma vez que, na área educacional, é provável que as dimensões estejam todas correlacionadas umas com as outras. Pesquisadores tais como Marsh (1991), Ramos (1997) e Val (1994) relatam que os fatores de seus respectivos instrumentos de avaliação da docência encontram-se correlacionados. Assim, faz-se necessário testar a hipótese de que os oito fatores do instrumento de avaliação da docência da UNISUL são todos correlacionados uns com os outros.

O modelo qualitativo teórico, apresentado no diagrama 4, que constitui a segunda hipótese a ser testada, sustenta a hipótese de que as oito dimensões do instrumentos são correlacionadas.

Hipótese 2: o modelo qualitativo teórico, apresentado no diagrama 4, é composto por oito dimensões correlacionadas confiáveis e válidas da qualidade do ensino no âmbito da sala de aula.

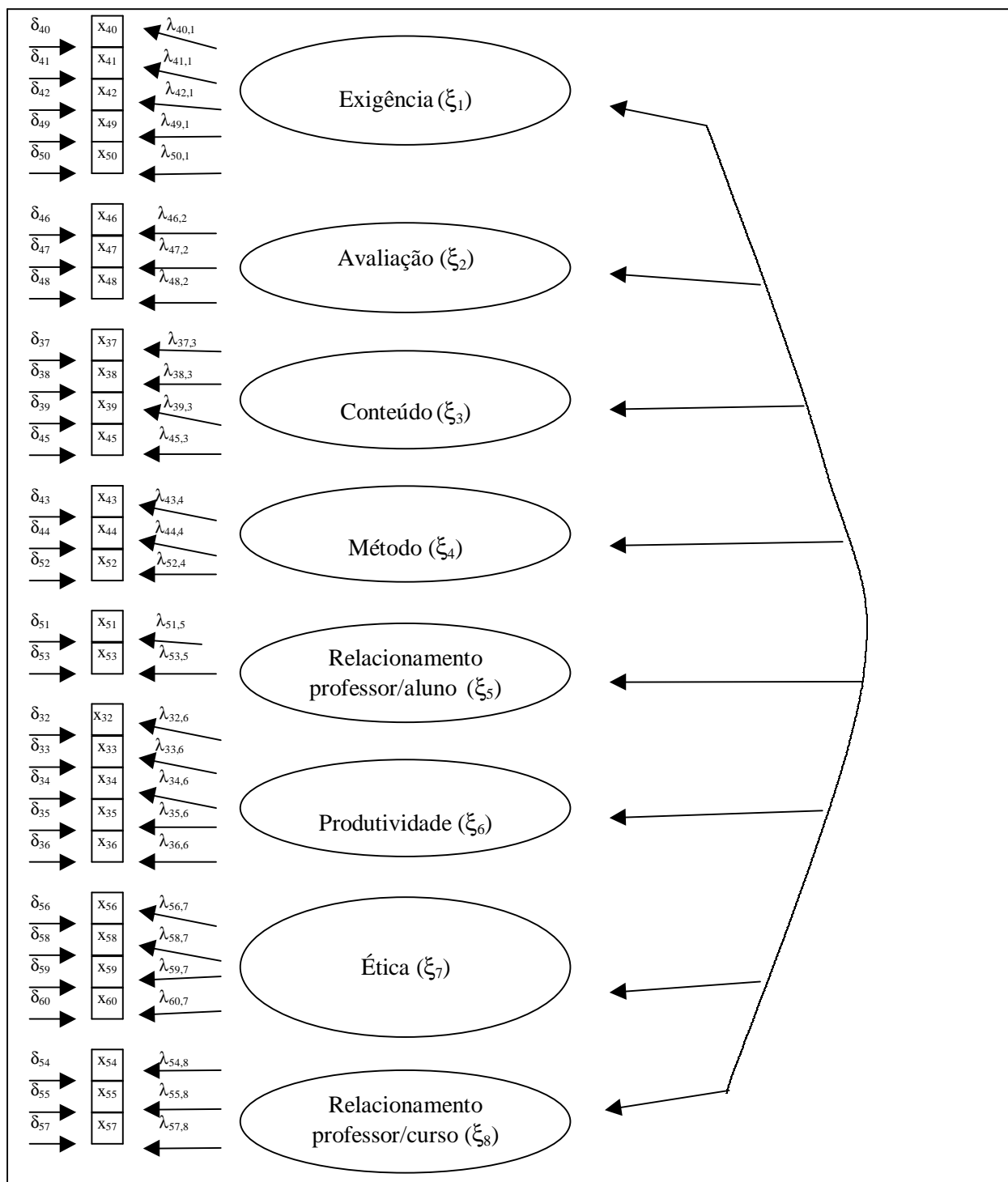


Diagrama 4: Diagrama de caminhos das características da qualidade docente (oito fatores latentes correlacionados).

As sub-hipóteses 2.1, 2.2, 2.3, 2.4, 2.5 2.6 enunciadas abaixo, fazem-se necessárias para a comprovação das propriedades fundamentais do instrumento de medida, ou seja, as propriedades de confiabilidade e de validade. A não rejeição dessas sub-hipóteses

conduz à aceitação da hipótese dois (2), evidencia-se o suporte empírico a favor das propriedades de confiabilidade e validade das dimensões operacionalizadas no instrumento, ou seja, a validade de construto.

Sub-hipótese 2.1: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre letivo de 2002, na UNISUL, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento de avaliação, ou seja, as oito dimensões correlacionadas do instrumento em análise.

Sub-hipótese 2.2: as dimensões correlacionadas, decorrentes das análises confirmatórias, são consistentes, conforme os critérios psicométricos.

Sub-hipótese 2.3: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre de 2002, no *campus* de Tubarão, constituem uma base de dados para o suporte empírico das oito dimensões correlacionadas do instrumento de medida.

Sub-hipótese 2.4: as dimensões correlacionadas, decorrentes das análises exploratórias e confirmatórias, para o total de disciplinas cursadas no *campus* de Araranguá, são consistentes, conforme os critérios psicométricos.

Sub-hipótese 2.5: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre de 2002, no *campus* da Pedra Branca, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento de medida, ou seja, das oito dimensões correlacionadas do instrumento de medida.

Sub-hipótese 2.6: as dimensões correlacionadas, decorrentes das análises confirmatórias, para o total de disciplinas cursadas no *campus* da Pedra Branca, são consistentes, conforme os critérios psicométricos.

Para a aceitação de uma das hipóteses 1 ou 2, é necessário que todos os parâmetros (λ) associados a cada flecha, apresentem significância estatística aceitável, coeficientes e erros de medida de acordo com os limites teóricos, bem como os índices da

qualidade de ajuste absoluto, incremental e de parcimônia relevantes, do ponto de vista estatístico.

A não rejeição das sub-hipóteses, enunciadas acima, são também indicativos da estabilidade da estrutura fatorial, ou seja, os alunos de *campus* distinto reconhecem da mesma forma os conceitos destinados a medirem a qualidade da instrução recebida intermediada pelo professor.

À medida que se vai aprofundando a análise dos dados, modelos alternativos surgirão, necessitando de pesquisa teórica e empírica para comprovar os novos modelos propostos. Caso existam fatores latentes altamente correlacionados, faz-se necessário modelar esses relacionamentos, por modelos fatoriais confirmatórios de 2ª ordem. Os modelos de segunda ordem também necessitam do apoio dos índices de qualidade de ajuste absoluto, incremental e de parcimônia, para evidenciar o grau de validade de construto.

Admitindo que o instrumento de medida utilizado pela UNISUL seja multidimensional, as hipóteses citadas abaixo julgam-se necessárias para o estabelecimento da validade de critério da escala. A validez de critério refere-se ao grau de eficácia com que se pode prever um indicador a partir da medição efetuada (SÁNCHEZ; SARAIBA, 1999). Quais sejam:

hipótese 3: as dimensões consistentes da escala possuem um efeito positivo e significativo, na valoração da qualidade global do curso por parte dos estudantes;

hipótese 4: as dimensões consistentes da escala possuem um efeito positivo e significativo, na valoração da variável critério “Indicaria este professor a outras turmas”.

hipótese 5: as dimensões consistentes da escala possuem um efeito positivo e significativo, na valoração da variável critério “O seu curso está exigindo o suficiente para sua formação profissional”.

As hipóteses 3, 4 e 5 serão aceitas, quando o coeficiente de correlação momento produto de Pearson entre as distintas dimensões do instrumento e as variáveis “Qualidade global do curso”, “indicaria este professor a outras turmas” e “o seu curso está exigindo o suficiente para sua formação profissional” são positivas e estatisticamente significativas. O estabelecimento da validade de critério indica que se pode prever um indicador a partir da medição efetuada.

Sendo os alunos clientes dos diversos serviços oferecidos pela UNISUL e que participam de forma ativa no processo de ensino e aprendizagem, existe a necessidade de investigar através dos modelos de regressão linear, quais dimensões os discentes consideram mais importante na valoração da qualidade total do curso ou da disciplina. A análise dos efeitos das dimensões latentes que integram a qualidade do ensino sobre a qualidade total do curso e das disciplinas será realizada através das seguintes hipóteses:

hipótese 6: cada uma das dimensões da escala possui um efeito positivo e significativo na valoração da qualidade da disciplina percebida pelos alunos.

hipótese 7: cada uma das dimensões da escala possui um efeito positivo e significativo, na valoração da qualidade total do curso percebida pelos alunos.

As hipóteses 6 e 7 serão aceitas quando os modelos de regressão linear múltipla evidenciam coeficientes betas estatisticamente significativos e todos os pressupostos das distribuições dos resíduos provenientes da modelagem satisfeitos.

O último objetivo a ser alcançado neste trabalho é a capacidade do instrumento para diagnosticar tipologias de qualidade docente em função da satisfação que revelam os

estudantes de graduação da UNISUL, e, desse modo, acumular mais evidências de validade para o instrumento em análise. Esse requisito é de fundamental importância para que possibilite um adequado uso das pontuações provenientes dos processos avaliativos.

Este estudo tem sido feito mediante análises agrupamentos K-médias. Como enfatizam Jornet e Suárez (1997, *apud* SUCH *et al.* 1999), esse tipo de análise pode ser utilizado para identificar tipologias de qualidade em avaliação de programas e/ou professores.

A análise de agrupamento é uma técnica multivariada que permite agrupar os professores em conglomerados, de maneira que os professores de um mesmo conglomerado sejam semelhantes entre si e bastante diferentes de um outro agrupamento de professores. A aplicação de metodologia de análises de agrupamentos, para identificar tipologias de qualidade docente, requer definir três questões básicas: medir as similitudes, formar os conglomerados e determinar o número de conglomerados.

Para identificar a solução de agrupamento mais representativa, Jornet e Suárez (1997, *apud* SUCH *et al.*, 1999) recomendam seguir os seguintes passos:

1. em igualdade de condições de diferenciação entre grupos, assume-se a solução de menor número de grupos;
2. entende-se que é mais representativa aquela tipologia que apresente maior número de diferenças estatisticamente significativas entre as variáveis originais,
3. Entende-se que uma solução é pouco representativa, quando na composição do grupo, alguns deles nascem de algum conglomerado pouco numeroso e, por isso, pode entender-se como marginal; e
4. entende-se que é preferível a escalabilidade total dos perfis às tipologias em que se produzam interações entre os grupos.

A capacidade do instrumento para diagnosticar tipologias de qualidade docente em função da satisfação que revelam os estudantes de graduação da UNISUL será detectada,

quando os perfis dos grupos de médias são paralelos e quando todas as soluções analisadas evidenciarem diferenças estatisticamente significativas entre os perfis dos grupos. A comprovação dessa hipótese, além de acumular evidências de validade, oportuniza o uso da média global para pontuar o docente (SUCH *et al.* 1999).

7.9 Limitações da pesquisa

Em geral, os professores não aderem passivamente à avaliação. A resistência manifestada decorre, segundo os professores, do fato de que as avaliações não refletem com segurança a realidade do ensino. Argumentam, ainda, que “Tamanho da Classe”, “Dificuldade da disciplina”, “Alunos poucos preparados para prosseguirem no curso elegido” são algumas das variáveis que provocam vieses negativos nos níveis de satisfação dos alunos quanto ao ensino intermediado pelo professor. Os professores ainda criticam o pouco uso que se faz das avaliações ocorridas em anos anteriores por parte das gerências dos cursos de graduação ou mesmo pelas instâncias superiores. Ponderam, ainda, que os resultados das avaliações não são amplamente discutidos com os alunos e que ações corretivas para o melhoramento do processo ensino-aprendizagem não chegam a ser concretamente operacionalizadas.

A pouca utilidade dos resultados da avaliação faz com que os alunos, quando convidados para avaliarem o seu desempenho na disciplina, o fazem de forma forçada, provocando, assim, uma influência não desejada nas avaliações dos docentes.

Todas essas variáveis, quando possível, necessitam ser analisadas e, se vier a se constituir em um viés, necessitam serem controladas durante o processo avaliativo.

Pesquisas americanas relatam que a motivação inicial do aluno (MARSH; DUNKIN, 1992); tipo de disciplina - obrigatória ou eletiva (ALEAMONI, 1981); notas esperadas (MARSH; DUNKIN, 1992); área acadêmica (FELDMAN, 1978; MARSH;

DUNKIN, 1992); não anonimato das avaliações (FELDMAN, 1979; MARSH; DUNKIN, 1992); objetivos da avaliação (FELDMAN, 1979; MARSH; DUNKIN, 1992), manifestam influências moderadas a baixas nas avaliações dos alunos. Pesquisas para medir o efeito dessas variáveis e outras necessitam serem investigadas localmente.

Quanto à validade do instrumento para diagnosticar tipologia de docentes, faz-se necessário um estudo longitudinal, em que as diversas disciplinas, ministradas pelo mesmo professor, poderão ser analisadas de maneira mais criteriosa.

7.10 Técnicas de análises de dados e software para análises

Conforme já discutidas nos capítulos 5 e 6, as técnicas multivariadas de análises fatorial exploratória e confirmatória serão utilizadas para comprovar as propriedades de confiabilidade e de validade do instrumento. O *software* utilizado foi o STATISTICA. For windows. O módulo *SEPATH*, desenvolvido por James Steiger, emprega a linguagem *Path 1* que permite a especificação do modelo em um formato muito legível e similar à construção de um diagrama de caminhos, conforme expostos pelos diagramas 3 e 4. Está também disponível o *Path Wizards* que permite guiar passo a passo mediante o processo de especificação para os tipos de modelos mais comuns.

8 ANÁLISE DAS PROPRIEDADES ESTATÍSTICA DE CONFIABILIDADE E DE VALIDADE DO INSTRUMENTO DA UNISUL “AVALIAÇÃO PELOS ALUNOS DO DESEMPENHO NA DISCIPLINA”

8.1 Introdução

O presente capítulo tem por objetivo evidenciar ou não as propriedades de confiabilidade e validade estatística do instrumento de avaliação de disciplina utilizada pela equipe de avaliação e clima organizacional da UNISUL.

Para testar as hipóteses: 1) o modelo qualitativo teórico apresentado no diagrama 3 é composto por oito dimensões não correlacionadas, consistentes e válidas da qualidade do ensino no âmbito da sala de aula; e 2) o modelo qualitativo teórico apresentado no diagrama 4 é composto por oito dimensões correlacionadas, consistentes e válidas da qualidade do ensino no âmbito da sala de aula, utilizou-se da abordagem de análise fatorial confirmatório apresentada no capítulo 6.

Os modelos qualitativos teóricos, apresentados nos diagramas 5 e 6, são os resultados de estudos teóricos realizados pela equipe a Avaliação Institucional e Clima Organizacional da UNISUL, portanto, a etapa de especificação do modelo que consiste no desenvolvimento de um modelo baseado na teoria, na construção de um diagrama de relações e na especificação do modelo de medida, *a priori*, já foi estabelecida.

8.2 Tratamento dos dados

Uma segura análise fatorial exploratória e confirmatória é realizada quando as distribuições das medidas utilizadas cumpram os critérios de homocedasticidade, linearidade e normalidade multivariada.

Os gráficos 1 e 2, exibem os *Box-plots* das medidas obtidos pelos indicadores medidos, decorrentes do processo avaliativo ocorrido na UNISUL, no segundo semestre letivo de 2002.

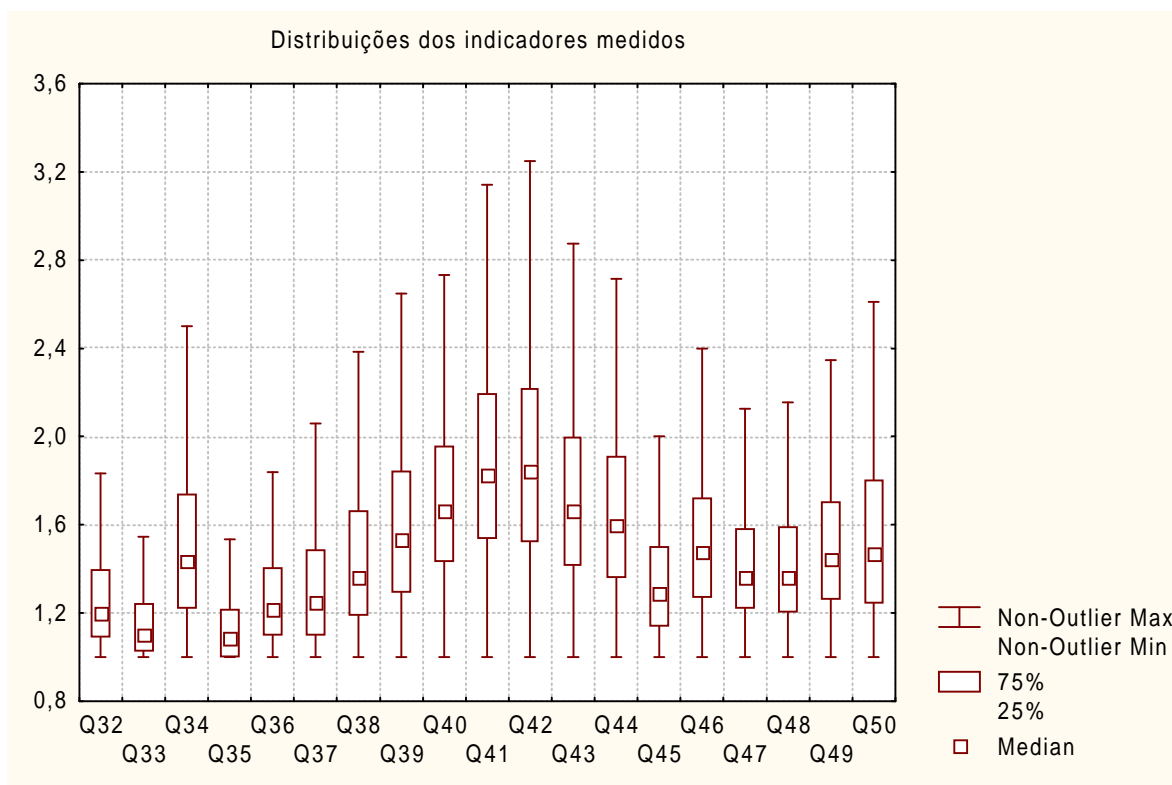


Gráfico 1: Distribuições dos indicadores avaliados. Geral UNISUL.

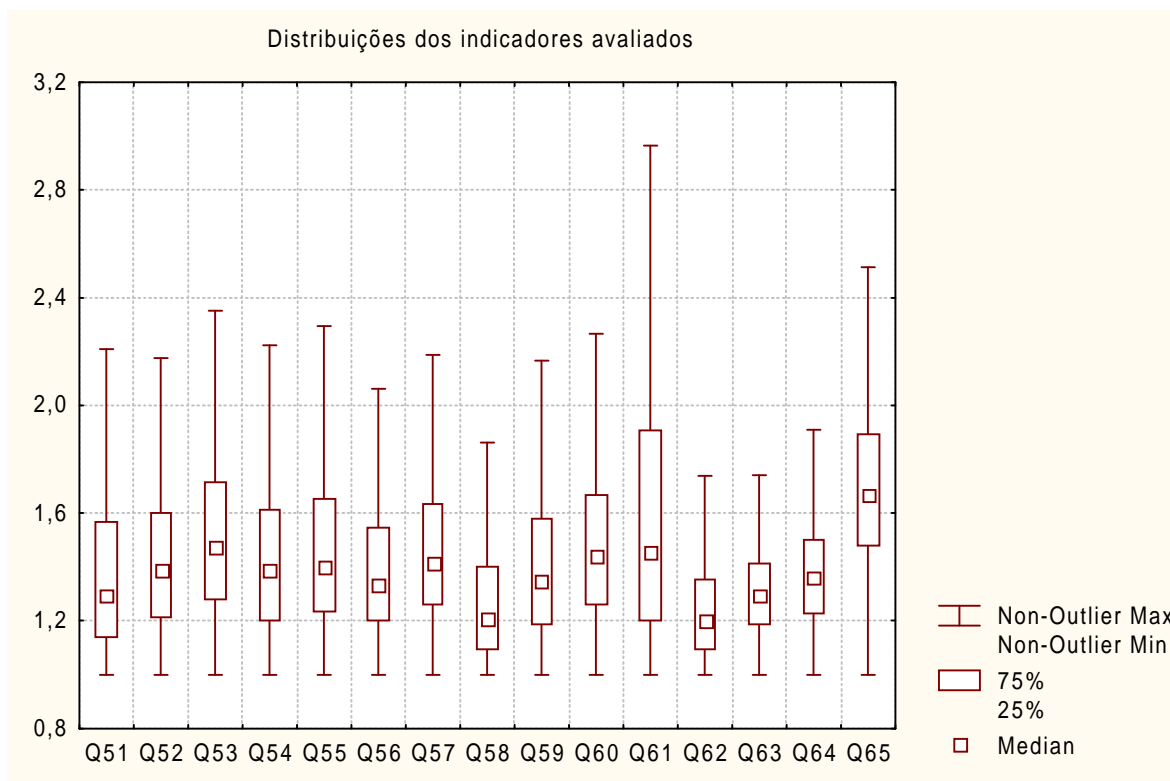


Gráfico 2: Distribuições dos indicadores avaliados. Geral UNISUL (continuação).

Quanto à variabilidade dos indicadores, os de número 32, 33, 35, 62 e 64 exibem baixa variabilidade, quando comparado com os demais. Indicadores que exibem correlação item-total e variabilidade baixa pouca contribuição trazem para a natureza da dimensão e, conseqüentemente, manifestarão baixa comunalidade nos estudos exploratórios. Os indicadores de números 34, 38, 39, 40, 41, 42, 43, 44, 50 e 61 são os que evidenciam maiores variabilidades, portanto, são os indicadores que revelaram maiores diferenças de opiniões por parte dos avaliadores. Observa-se que os indicadores de números 39, 40, 41, 42, 43, e 44, além de evidenciarem altas variabilidades, são também os que apresentaram os piores desempenhos medianos. Assim, conclui-se que os indicadores exibidos pelos gráficos 1 e 2 exibem variabilidades bastante distintas sendo, portanto, heterocedástica.

Quanto à normalidade dos dados empíricos, o coeficiente de normalidade multivariada de Nardia (1970) é igual a 0,450. Desta forma, afasta-se muito do padrão esperado, que é zero. Essas informações conduzem para a transformação nos dados empíricos.

Espera-se que a transformação dos dados empíricos, além de diminuir o excesso de assimetria e de curtose desses indicadores, tornem-se razoavelmente homocedástico.

Submetendo os dados empíricos ao operador $1/\mathbf{x}$, a distribuição multivariada dos dados transformados aproxima-se razoavelmente da distribuição multinormal. O coeficiente de normalidade multivariada de Nardia passou para 0,180, ainda alto, quando comparado com o padrão. Em decorrência dos resultados da transformação, os dados transformados são os que serão usados para testar os modelos qualitativos teóricos apresentados nos diagramas 3 e 4 e outros.

8.3 Estimativas da confiabilidade dos indicadores e das dimensões do instrumento de avaliação de disciplinas da UNISUL

Durante a fase de especificação do modelo de medida de indicadores exógenos, além de determinar o número de indicadores de cada dimensão, deve-se considerar a confiabilidade dos indicadores e dos construtos, para garantir que os modelos teóricos sejam testados a partir de medidas isentas de erros aleatórios.

A tabela 13 apresenta as dimensões dos modelos de medida especificados nos diagramas 3 e 4 com os seus respectivos indicadores, a confiabilidade de cada indicador expressa pela correlação item-total, e a confiabilidade da dimensão medido pelo coeficiente α de Cronbach para o total de disciplinas avaliadas na UNISUL, no segundo semestre letivo de 2002.

Tabela 13: Confiabilidade dos indicadores e das dimensões do instrumento de avaliação de disciplinas da UNISUL – Geral.

Dimensão/indicador	Correlação Item-total	Confiabilidade da dimensão
Exigência		
40	0,793	
41	0,670	
42	0,707	0,8851
49	0,734	
50	0,725	
Avaliação		
46	0,840	
47	0,812	0,9184
48	0,855	
Conteúdo		
37	0,793	
38	0,818	0,914
39	0,849	
45	0,754	
Método		
43	0,770	
44	0,789	0,884
52	0,767	
Relacionamento professor/ aluno		
51	0,794	
53	0,836	0,867
Produtividade		
32	0,671	
33	0,633	
34	0,642	0,844
35	0,546	
36	0,771	
Ética		
56	0,746	
58	0,770	
59	0,829	0,909
60	0,835	
Relação professor/curso		
54	0,838	
55	0,770	0,891
57	0,755	

As correlações item-total estratificados, segundo o *campus*, apresentam um padrão semelhante como as apresentadas na tabela 13. A confiabilidade das dimensões estratificadas por *campus* segue apresentadas na tabela 14.

Tabela 14: Confiabilidades das dimensões do instrumento de avaliação de disciplinas da UNISUL, estratificadas segundo o campus.

Dimensão	Campus			Geral da UNISUL
	Tubarão	Araranguá	Pedra Branca	
Exigência	0,878	0,880	0,892	0,885
Avaliação	0,915	0,923	0,922	0,918
Conteúdo	0,912	0,919	0,918	0,914
Método	0,893	0,895	0,886	0,884
Relacionamento professor aluno	0,836	0,869	0,890	0,867
Produtividade	0,800	0,834	0,838	0,844
Ética	0,898	0,924	0,919	0,909
Relacionamento professor curso	0,890	0,923	0,890	0,891

Ao se considerar que a confiabilidade do indicador deve superar a 0,5 (HAIR *et al.*, 1999); e a confiabilidade da dimensão deve ser maior que 0,7 (NUNNALLY, 1987), os resultados reportados na tabela 13 e 14 indicam que todos os indicadores e dimensões são consistentes. Esses resultados confirmam que as oito dimensões do instrumento de avaliação de disciplina da UNISUL são consistentemente avaliadas pelos discentes e indicam que os modelos teóricos especificados nos diagramas 3 e 4 são testados a partir de medidas com erros aleatórios minimizados.

8.4 Identificação do modelo com oito fatores não correlacionados e do modelo com oito fatores correlacionados e tamanhos de amostras.

Depois de especificado que os modelos de medida exibidos nos diagramas 3 e 4 podem ser estimados, utilizando-se de mediadas confiáveis, avalia-se, a seguir, a capacidade

dos modelos para estimar os parâmetros únicos dos modelos de medida, ou seja, avalia-se o seu nível de identificação.

A tabela 15 relata os números de parâmetros e de correlações a serem estimados, totais de correlações utilizadas na modelagem e o número de graus de liberdade necessários para estimar os modelos de medida de indicadores exógenos apresentados nos diagramas 3 e 4.

Tabela 15: Graus de liberdade dos modelos de medida apresentado nos diagramas 3 e 4.

Modelos	Nº de parâmetros	Correlações entre fatores	Total de correlação	Graus de liberdade
Figura 9.1	26	0	325	299
Figura 9.2	26	28	325	271

Como os modelos de medidas, citados acima, apresentam um número positivo de graus de liberdade, é possível estimar os dois modelos de medida com indicadores de ajuste corretos com o maior grau de liberdade possível.

Na determinação do tamanho das amostras é levado em consideração a sensibilidade do método de estimação (máxima verossimilhança) para grandes amostras, o tamanho da população e o fato de que os dados empíricos, mesmo transformados, não apresentem coeficiente de normalidade multivariada de Nardia igual a zero. Segundo recomendações de Hair *et al.* (1999), a falta de normalidade multivariada dos dados empíricos eleva de 10 para 15 o número de avaliadores por parâmetro a ser estimado.

A tabela 16 apresenta o total de disciplinas avaliadas no geral da UNISUL e estratificadas por *campus* bem como o tamanho das amostras utilizadas para a estimação das correlações a ser usada nas estimativas dos parâmetros dos modelos.

Tabela 16: Número de disciplinas utilizadas para comprovação dos modelos, expostos nos fluxogramas 5 e 6.

UNISUL	População	Tamanho da amostra (modelo com 8 dimensões latentes não correlacionadas)	Tamanho da amostra (modelo com 8 dimensões latentes correlacionadas)
Tubarão	956	675	675
Araranguá	249	249	249
Pedra Branca	837	675	675
Geral	2042	675	675

8.5 Estimação do modelo com oito fatores latentes não correlacionados e do modelo com oito fatores latentes correlacionados

O uso da matriz de correlação para estimar os modelos de medidas possibilita comparações diretas dentro dos modelos. As matrizes de correlações usadas nas estimações dos modelos com oito fatores latentes não correlacionadas e dos modelos com oito fatores latentes correlacionadas estão apresentados nos anexos VI, VII, VIII e IX.

As estimativas dos parâmetros desses modelos são obtidas pelo método de estimação de máxima verossimilhança (MV). Decorrem das estimações desses modelos, os índices de ajustes absolutos, incrementais e de parcimônia, que são utilizados conjuntamente para comparar e avaliar os modelos estimados. A tabela 17 apresenta as medidas de adequação do ajuste para os dois modelos, considerados os dados globalizados da UNISUL.

Tabela 17: 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

χ^2 : Estatística Qui-quadrado ($p_{\text{valor}} > 0,05$), RMSR: Erro de aproximação quadrático médio ($< 0,08$), GFI: Índice de bondade de ajuste ($> 0,90$), PNC: Parâmetro de não centralidade (menor melhor), PNCE: Parâmetro de não centralidade da escala (menor melhor), AGFI: Índice de adequação do ajuste corrigido ($> 0,90$), TLI: Índice de ajuste de Tucker-Lewis ($> 0,90$), BBI: Índice de ajuste normalizado ($> 0,90$), PNFI: Índice de ajuste normalizado de parcimônia (maior melhor), PGFI: Índice de bondade de ajuste (maior melhor), AIC: Critério de informação de Akaike (menor melhor).

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 compara-se o modelo com oito fatores correlacionados com o modelo com fatores independentes. Ressalta-se que, sendo o modelo com oito fatores correlacionados superiores aos modelo com fatores independentes, este apresenta erro de aproximação quadrático médio muito acima do recomendado; ou seja, acima de 0,08. Os outros 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 para o modelo com oito fatores correlacionados, indicando que este apresenta maior poder de explicação, quando comparado com o modelo de oito fatores independentes, mas estes precisam ser melhorados para alcançarem melhor poder de explicação.

Os índices de ajuste de parcimônia são os que apresentam, mais significados dentro de uma estratégia de modelização competitiva, por oferecer medidas do ajuste do modelo por coeficiente estimado. Os valores dos índices PNFI, PGFI e o AIC confirmam o modelo com oito fatores correlacionados, como superior ao modelo com oito fatores independentes.

Os valores dos parâmetros λ e o seu respectivo erro padrão, o erro de medida δ dos indicadores decorrentes do processo de estimação por máxima verossimilhança do modelo com oito fatores correlacionados, estimado para o geral da UNISUL, estão apresentados na tabela 18.

Tabela 18: Valores dos coeficientes lâmbda (λ), erro padrão e erro de medida dos indicadores – Geral da UNISUL

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

* Os indicadores de números 32, 33 e 35 foram excluídos da análise em função de evidenciarem baixas variabilidades, conseqüentemente, evidenciam pouca contribuição na formação desse construto.

** Todos os lâmbdas (λ) são estatisticamente significativo ao nível de 1%.

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

A qualidade do modelo estimado é também examinada através da análise dos resíduos padronizados. O ideal é que esses resíduos padronizados sejam próximos de; o que indicaria uma reprodução próxima da matriz de correlação observada. Tomando como nível de significância de 0,01, aqueles resíduos padronizados com valores superiores $\pm 2,58$ indicariam uma diferença significativa entre a matriz de correlação observada e a matriz de correlação estimada. O gráfico 3 apresenta o gráfico de probabilidade normal para os resíduos do modelo com oito fatores correlacionados, geral UNISUL.

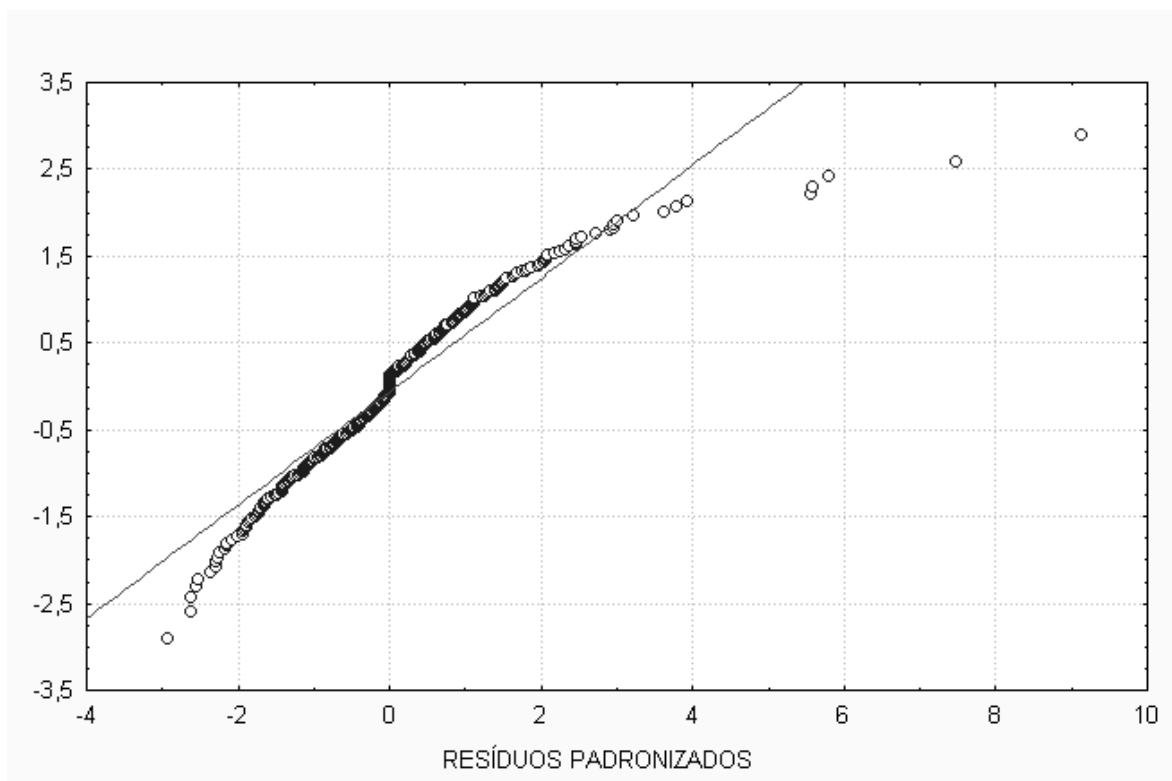


Gráfico 3: Gráfico de probabilidade normal dos resíduos do modelo com oito fatores correlacionados, geral UNISUL.

O gráfico 3 exibe alguns resíduos padronizados que superam $\pm 2,58$, sugerindo que o modelo proposto deva ser revisado.

As medidas de adequação do ajuste para o modelo de medida com oito fatores não correlacionados e do modelo com oito fatores correlacionados, estratificados por *campus*, estão apresentadas na tabela 19.

Tabela 19: Índices de adequação do ajuste dos modelos com oito dimensões ou fatores não correlacionados e oito dimensões ou fatores correlacionados, estratificados segundo o campus.

Medidas de adequação do ajuste	Campus Tubarão		Campus Araranguá		Campus Pedra Branca	
	Modelo com 8 fatores não correlacionadas	Modelo com 8 fatores correlacionadas	Modelo com 8 fatores não correlacionadas	Modelo com 8 fatores correlacionadas	Modelo com 8 fatores não correlacionadas	Modelo com 8 fatores correlacionadas
Absoluto de ajuste						
χ^2	10523,40 (gl 301)	3951,11 (gl 271)	4101,99 (gl 301)	1411,26 (gl 271)	10546,70 (gl 301)	3331,64 (271)
RMSR	0,267	0,16	0,270	0,146	0,270	0,142
GFI	0,373	0,639	0,360	0,654	0,368	0,688
PNC	10222,4	3680,11	3800,99	1140,26	10245,70	3060,64
PNCE	15,144	5,452	15,265	4,579	15,179	4,534
Incrementais de ajuste						
AGFI	0,269	0,532	0,254	0,551	0,263	0,596
TLI	0,456	0,783	0,485	0,828	0,479	0,827
BBI	0,490	0,808	0,505	0,830	0,511	0,846
Ajuste de parcimônia						
PNFI	0,454	0,674	0,468	0,692	0,473	0,705
PGFI	0,346	0,533	0,333	0,545	0,341	0,574
χ^2 (Normalizado)	34,961	14,579	13,628	5,208	35,039	12,294
AIC	15,762	6,100	16,944	6,336	15,796	5,180

χ^2 : Estatística Qui-quadrado ($p_{\text{valor}} > 0,05$), RMSR: Erro de aproximação quadrático médio ($< 0,08$), GFI: Índice de adequação do ajuste ($> 0,90$), PNC: Parâmetro de não centralidade (menor melhor), PNCE: Parâmetro de não centralidade da escala (menor melhor), AGFI: Índice de adequação do ajuste revisado ($> 0,90$), TLI: Índice de ajuste de Tucker-Lewis ($> 0,90$), BBI: Índice de ajuste normalizado ($> 0,90$), PNFI: Índice de ajuste normalizado de parcimônia (maior melhor), PGFI: Índice de bondade de ajuste (maior melhor), AIC: Critério de informação de Akaike (menor melhor).

Embora os modelos tenham sido estimados com tamanhos de amostras diferentes, os índices de ajustes do modelo com oito fatores latentes independentes são muito semelhantes. O mesmo acontece com os modelos com oito fatores latentes correlacionados, que evidenciam os melhores índices de adequação do ajuste.

A superioridade do modelo com oito fatores correlacionados, evidenciados pelos índices de ajustes absolutos, incrementais e de parcimônia, conduz à rejeição do modelo qualitativo teórico apresentado no diagrama 3 e, conseqüentemente, de todas as suas sub-hipóteses. Essa conclusão não garante que o modelo qualitativo teórico apresentado no diagrama 4 e suas sub-hipóteses sejam aceitas plenamente.

Quanto à correlação entre os fatores latentes, a tabela 20 apresenta as correlações entre os fatores latentes dos modelos apresentados nas tabelas 17 e 19.

Tabela 20: Correlações entre os fatores latentes.

Correlação	Geral UNISUL	Tubarão	Araranguá	Pedra Branca
2*1	0,819	0,769	0,826	0,847
3*1	0,977	0,981	0,968	0,973
4*1	0,951	0,936	0,933	0,963
5*1	0,789	0,757	0,780	0,781
6*1	0,935	0,965	0,919	0,942
7*1	0,863	0,849	0,878	0,893
8*1	0,872	0,852	0,890	0,892
3*2	0,842	0,793	0,859	0,868
4*2	0,855	0,815	0,906	0,866
5*2	0,833	0,798	0,814	0,813
6*2	0,734	0,731	0,704	0,747
7*2	0,853	0,831	0,902	0,865
8*2	0,759	0,704	0,819	0,788
4*3	0,971	0,961	0,968	0,986
5*3	0,870	0,824	0,887	0,864
6*3	0,904	0,933	0,867	0,882
7*3	0,895	0,873	0,923	0,931
8*3	0,827	0,802	0,850	0,880
5*4	0,933	0,925	0,951	0,930
6*4	0,823	0,843	0,769	0,836
7*4	0,840	0,944	0,993	0,963
8*4	0,889	0,839	0,887	0,908
6*5	0,667	0,679	0,662	0,657
7*5	0,991	0,992	0,996	0,981
8*5	0,826	0,764	0,800	0,838
7*6	0,746	0,774	0,743	0,776
8*6	0,764	0,774	0,756	0,783
8*7	0,885	0,837	0,910	0,913

As correlações entre fatores latentes consideradas elevadas, geralmente acima de 0,9, significa que tais fatores latentes estão explicando informação redundante, então não representam construtos diferentes. Assim, de acordo com esse critério, os fatores 3*1, 4*1, 6*1, 4*3, 6*3, 7*3, 5*4, 7*4, 7*5, reportados na tabela 20, não evidenciam que esses construtos sejam diferentes. Portanto, os baixos desempenhos apresentados pelos modelos estimados com oito fatores correlacionados são decorrentes da má especificação do modelo e/ou das altas correlações entre os fatores latentes.

Quanto à confiabilidade composta e à variância extraída desses fatores latentes, todos os resultados superaram os limites mínimos de aceitação, ou seja, 0,7 para a

confiabilidade composta e 0,5 para a variância extraída, conforme pode ser visualizado na tabela 21. Devido à similaridade desses resultados, quando estratificados por *campus*, é apresentado somente os índices de confiabilidades do modelo geral UNISUL.

Tabela 21: Confiabilidade composta e variância extraída dos fatores latentes correlacionados - geral UNISUL.

Fatores latentes	Confiabilidade composta	Variância extraída
Exigência	0,865	0,572
Avaliação	0,922	0,832
Conteúdo	0,910	0,716
Método	0,822	0,726
Relacionamento professor aluno	0,857	0,749
Produtividade	0,838	0,724
Ética	0,918	0,737
Relação professor curso	0,889	0,728

A confiabilidade composta indica a consistência interna com que todos os indicadores estão medindo o construto, enquanto que a variância extraída revela a quantidade global da variância dos indicadores explicados pela variável latente. Os índices de confiabilidade composta e variância extraída, reportada na tabela 21 indicam quão consistentemente os discentes valoram os indicadores do instrumento de avaliação.

Apesar de todos os indicadores de confiabilidade assegurarem que os modelos foram estimados, utilizando-se de medidas confiáveis, é necessário que o modelo de medida com oito fatores latentes correlacionados correspondente ao diagrama 4 seja reavaliado, até obter índices de qualidade do ajuste maiores. Assim, possibilita dotar o instrumento em análise de maior grau de validade de construto. Esta conclusão decorre dos seguintes fatos não excludentes:

1. os índices de adequação do ajuste absoluto, incremental e de parcimônia não são satisfatórios sob o ponto de vista estatístico;
2. existem indicadores que apresentam elevados erros de medida;
3. há fatores latentes que são altamente correlacionados, portanto, um deles pode ser removidos; e
4. uma quantidade de resíduos que excedem ao valor de corte resultante do nível de significância escolhido.

8.6 Modificação do modelo

O estudo das novas dimensões que configuram a qualidade do ensino, no âmbito da sala de aula, é realizado mediante a aplicação dos dados empíricos à análise fatorial exploratória. No primeiro momento, é realizado um estudo da adequação dos dados à análise fatorial exploratória, seguido da extração dos fatores necessários para restringir o número de indicadores da escala através da técnica de componentes principais. Segue com a rotação ortogonal e oblíqua. Posteriormente, é realizada a interpretação dos fatores.

A análise da matriz de correlação entre todos os indicadores da escala permite confirmar que os indicadores estão correlacionados e que suas características são adequadas para realizar uma análise fatorial exploratória. Essa e outras medidas do grau de associação entre os indicadores são apresentadas na tabela 22. Os resultados justificam o emprego da análise fatorial exploratória.

Tabela 22: Medidas do grau de associação entre os indicadores* .

Medidas	Geral da UNISUL	Tubarão	Araranguá	Pedra Branca
Matriz de correlações	Indicadores correlacionados	Indicadores correlacionados	Indicadores correlacionados	Indicadores correlacionados
Determinante da matriz de correlação	-13,109	-12,874	-14,567	-13,376
Teste de esfericidade de Bartlett	26630,93 (gl 325)	12172,37 (gl 325)	3474,23 (gl 325)	11055,26 (gl 325)
Índice de Kaiser-Meyer-Olkin	P(valor) = 0,00000 0,947	P(valor) = 0,00000 0,940	P(valor) = 0,00000 0,944	P(valor) = 0,00000 0,949

* As matrizes de correlações utilizadas nos estudos exploratórios, estão apresentadas nos anexos II, III, IV e V.

Os números de fatores a serem retidos foram evidenciados pelo critério de contraste de caída, o qual se adaptou melhor aos dados. A seguir, diversas rotações foram realizadas para clarificar o significado do fator, sendo o procedimento de rotação ortogonal **VARIMAX** o que melhor evidenciou a estrutura dos fatores. Posteriormente, foi realizado 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, ortogonal e oblíqua.

As tabelas 23, 24, 25 e 26 apresentam a composição dos fatores, a comunalidade de cada indicador, a matriz de correlação entre os fatores e confiabilidade desses, o autovalor, a porcentagem da variância explicada e a porcentagem de variância explicada, acumulada para o total de disciplinas avaliadas na UNISUL e estratificadas, segundo o *campus*.

Tabela 23: Matriz fatorial rotacionada – geral UNISUL.

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	

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

Tabela 24: Matriz fatorial rotacionada – Campus Tubarão.

Indicadores	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidade
X ₃₄	0,8575				0,7827
X ₃₆	0,6587				0,5798
X ₃₇	0,8192				0,7364
X ₃₉	0,7856				0,8122
X ₄₉	0,8293				0,8168
X ₅₀	0,8075				0,8898
X ₅₁		0,8037			0,7958
X ₅₂		0,6987			0,7276
X ₅₃		0,7974			0,7655
X ₅₈		0,8258			0,7181
X ₆₀		0,7982			0,7765
X ₄₀			0,7119		0,6796
X ₄₁			0,8480		0,5857
X ₄₂			0,8676		0,6729
X ₄₆				0,7257	0,7291
X ₄₇				0,7681	0,6846
X ₄₈				0,7148	0,7923
Correlações entre fatores					
	Fator 1	Fator 2	Fator 3	Fator 4	
Fator 1	0,9487				
Fator 2	0,6538	0,9312			
Fator 3	0,5861	0,5101	0,8654		
Fator 4	0,7081	0,7500	0,4552	0,9138	
Autovalor	9,900	1,7537	1,4313	0,7691	
% Variância explicada	58,7644	10,3161	8,4194	4,5247	
% Variância explicada acumulada	58,7644	69,0805	77,500	82,0247	

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

Tabela 25: Matriz fatorial rotacionada – Campus Araranguá.

Indicadores	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidade
X ₃₄	0,8485				0,7894
X ₃₆	0,6674				0,5846
X ₃₇	0,8399				0,8117
X ₃₉	0,7361				0,8322
X ₄₉	0,8025				0,8146
X ₅₀	0,7990				0,8663
X ₅₁		0,8253			0,8407
X ₅₂		0,6568			0,7734
X ₅₃		0,7897			0,7837
X ₅₈		0,7918			0,8168
X ₆₀		0,7577			0,8487
X ₄₀			0,7755		0,7224
X ₄₁			0,7871		0,5782
X ₄₂			0,8539		0,6563
X ₄₆				0,7664	0,7782
X ₄₇				0,7126	0,7776
X ₄₈				0,6957	0,8310
Correlações entre fatores					
	Fator 1	Fator 2	Fator 3	Fator 4	
Fator 1	0,9497				
Fator 2	0,7194	0,9493			
Fator 3	0,6257	0,6159	0,8717		
Fator 4	0,8053	0,7476	0,5505	0,9300	
Autovalor	10,8953	1,3810	1,2887	0,6412	
% Variância explicada	64,0898	8,1232	7,5806	3,7717	
% Variância explicada acumulada	64,0898	72,2130	79,7936	83,5653	

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade do fator.

Tabela 26: Matriz fatorial rotacionada – Campus Pedra Branca.

Indicadores	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidade
X ₃₄	0,8050				0,8036
X ₃₆	0,7411				0,5887
X ₃₇	0,7897				0,7398
X ₃₉	0,6732				0,8111
X ₄₉	0,7395				0,8016
X ₅₀	0,7255				0,8771
X ₅₁		0,8156			0,8282
X ₅₂		0,7198			0,7344
X ₅₃		0,8319			0,7957
X ₅₈		0,8156			0,7636
X ₆₀		0,7973			0,8072
X ₄₀			0,7392		0,6970
X ₄₁			0,8330		0,6108
X ₄₂			0,8742		0,6922
X ₄₆				0,7759	0,7470
X ₄₇				0,6947	0,7189
X ₄₈				0,6954	0,7827
Correlações entre fatores					
	Fator 1	Fator 2	Fator 3	Fator 4	
Fator 1	0,9556				
Fator 2	0,5727	0,9494			
Fator 3	0,7268	0,6708	0,8725		
Fator 4	0,7724	0,5413	0,7858	0,9300	
Autovalor	64,2008	8,6814	6,8473	3,6034	
% Variância explicada	64,2008	5,6814	6,8473	3,6034	
% Variância explicada acumulada	64,2008	72,8822	79,7295	83,3330	

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade do fator.

Os indicadores X₃₈, X₄₃, X₄₄, X₄₅, X₅₄, X₅₅, X₅₆, X₅₇ e X₅₉ foram removidos da análise, em todos os estudos exploratórios, por não apresentarem carga fatorial suficiente, ou por não discriminarem bem entre os fatores.

Determinado o número de fatores que simplifique a escala, segue a análise da confiabilidade e da validade dos indicadores, assim como da confiabilidade de cada fator retido pela análise fatorial exploratória. A confiabilidade de cada indicador é constituída pela variância comum e pela variância específica, enquanto que a sua validade é constituída

somente pela variância comum. A maioria dos autores que recorreram à análise fatorial exploratória para estudar os construtos existentes na estrutura de seus instrumentos de avaliação, recomenda 0,5 como ponto de corte. A tabela abaixo apresenta os índices de confiabilidade de cada indicador, para o geral da UNISUL e estratificado segundo o *campus*.

Tabela 27: Confiabilidade dos indicadores decorrentes dos estudos exploratórios. Geral da UNISUL e estratificado segundo o Campus.

Indicador	Geral da UNISUL	Campus de Tubarão	Campus de Araranguá	Campus de Pedra Branca
X ₃₄	0,7022	0,7353	0,7200	0,6480
X ₃₆	0,4646	0,4339	0,4454	0,5492
X ₃₇	0,6639	0,6711	0,7054	0,6236
X ₃₉	0,5543	0,6172	0,5418	0,4532
X ₄₉	0,6306	0,6877	0,6440	0,5469
X ₅₀	0,6058	0,6521	0,6384	0,5264
X ₅₁	0,6636	0,6459	0,6811	0,6652
X ₅₂	0,4904	0,4882	0,4314	0,5181
X ₅₃	0,6579	0,6358	0,6236	0,6921
X ₅₈	0,6777	0,6819	0,6269	0,6652
X ₆₀	0,6268	0,6371	0,5741	0,6357
X ₄₀	0,5282	0,5068	0,6014	0,5464
X ₄₁	0,7024	0,7191	0,6195	0,6939
X ₄₂	0,7533	0,7527	0,7291	0,7642
X ₄₆	0,5515	0,5266	0,5874	0,6020
X ₄₇	0,5194	0,5900	0,5078	0,4826
X ₄₈	0,4949	0,5109	0,4840	0,4836

A maioria dos indicadores apresenta confiabilidades bastante elevadas, o que indica que explicam uma alta porcentagem da variância dos fatores latentes. Os indicadores que apresentam confiabilidade abaixo de 0,5 foram mantidos na escala para assegurar a confiabilidade do construto. Esses indicadores serão removidos da análise se os estudos confirmatórios evidenciarem erros de medida superiores a 0,5.

Quanto à confiabilidade dos fatores, medidos pelo *Alfa de Cronbach*, apresentados em negrito, na diagonal principal da matriz de correlações entre os fatores retidos pela análise fatorial exploratória, em todos os modelos exploratórios superou a cota mínima estabelecida por Nunnally (1987), ou seja, 0,7. Assim, pode-se concluir que todos os fatores foram medidos com quantidades de erros aleatórios admissíveis.

As saturações ou correlações evidenciadas entre o indicador e o fator são relevantes sob o ponto de vista estatístico, a maioria excedem a 0,7 e são unidimensionais. Quanto à validade desses indicadores, todos os indicadores, em todos os modelos exploratórios, exibem comunalidade superiores a 0,5, mostrando que todos os indicadores foram suficientemente explicados pelos fatores decorrentes do processo de análise fatorial exploratória.

As quatro estruturas fatoriais obtidas sustentam a hipótese de estabilidade fatorial e dão ênfase aos seguintes aspectos do processo de ensino e de aprendizagem, no âmbito da sala de aula:

- FATOR 1: reconhece que o professor que administra o tempo em sala de aula de forma produtiva, que está desenvolvendo o plano de ensino, que revela domínio de conteúdo na sua disciplina, que apresenta uma bagagem intelectual estimulante para os alunos, que trabalha a disciplina com um nível de exigência suficiente e que conduz o ensino da disciplina com qualidade são os atributos mais importantes para uma docência com qualidade. Essa conclusão decorre do fato de que o primeiro fator retido pela análise fatorial é o que carrega a maior porcentagem da variância explicada. Conforme as tabelas 23, 24, 25 e 26, o primeiro fator explica acima de 60% da variabilidade total dos dados em todos os casos. Engloba cinco indicadores importantes 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 teoricamente relacionado positivamente com os demais fatores.

- FATOR 2: 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 apresentados no capítulo 2.

- FATOR 3: 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 a leitura de livros, textos, jornais e revistas complementares às aulas. Essas ações do professor são necessárias para manter os alunos motivados e, conseqüentemente, obter um aprendizado mais efetivo. Esse fator está teoricamente correlacionado com o primeiro e o segundo fator e é denominado de “motivação dos discentes”.

- FATOR 4: os seguintes indicadores estão relacionados com esse fator: 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 denominado de “avaliação” está presente em todos os instrumentos de avaliação da docência, apresentados no capítulo II. Esse fator correlaciona-se com os demais.

Portanto, pode-se concluir que o instrumento de avaliação da docência pelos alunos que, *a priori*, foi concebido com oito dimensões, pode ser adequadamente explicado

por quatro dimensões ou fatores correlacionadas que explicam aproximadamente 83% da variabilidade dos indicadores retidos para análise. São elas: competência, relacionamento e ética, motivação, e avaliação. Os indicadores dessas dimensões são unidimensionais, evidenciando a sua validade discriminante.

Tendo assegurado que a escala é adequada às suas definições conceituais, que os indicadores são unidimensionais e que alcança níveis suficientes de confiabilidade, faz-se necessária uma avaliação final, ou seja, a validação da escala. Essa etapa é imprescindível, uma vez que a análise fatorial exploratória é utilizada para ajudar o pesquisador com conhecimentos teóricos insuficientes, na formulação ou reformulação dos conceitos.

Após a reformulação dos conceitos, faz-se necessária à comprovação da legitimidade desses construtos. Essa necessidade decorre dos seguintes fatos: 1) o primeiro fator obtido pela técnica de componentes principais apresenta todos os indicadores intensamente saturados sobre ele; e 2) num segundo momento, é feita a rotação dos eixos, possibilitando que distintos pesquisadores utilizem técnica de rotação diferente e obtenha fatores latentes que possibilite interpretações distintas para um mesmo instrumento de avaliação.

Do exposto, duas novas hipóteses são testadas para poder dotar o instrumento em apreciação de maior grau de validade de construto. Quais sejam: as hipóteses oito e nove, apresentadas a seguir.

Hipótese 8: o modelo qualitativo teórico, apresentado no diagrama 5 é composto por um fator latente confiável e válido da qualidade do ensino no âmbito da sala de aula.

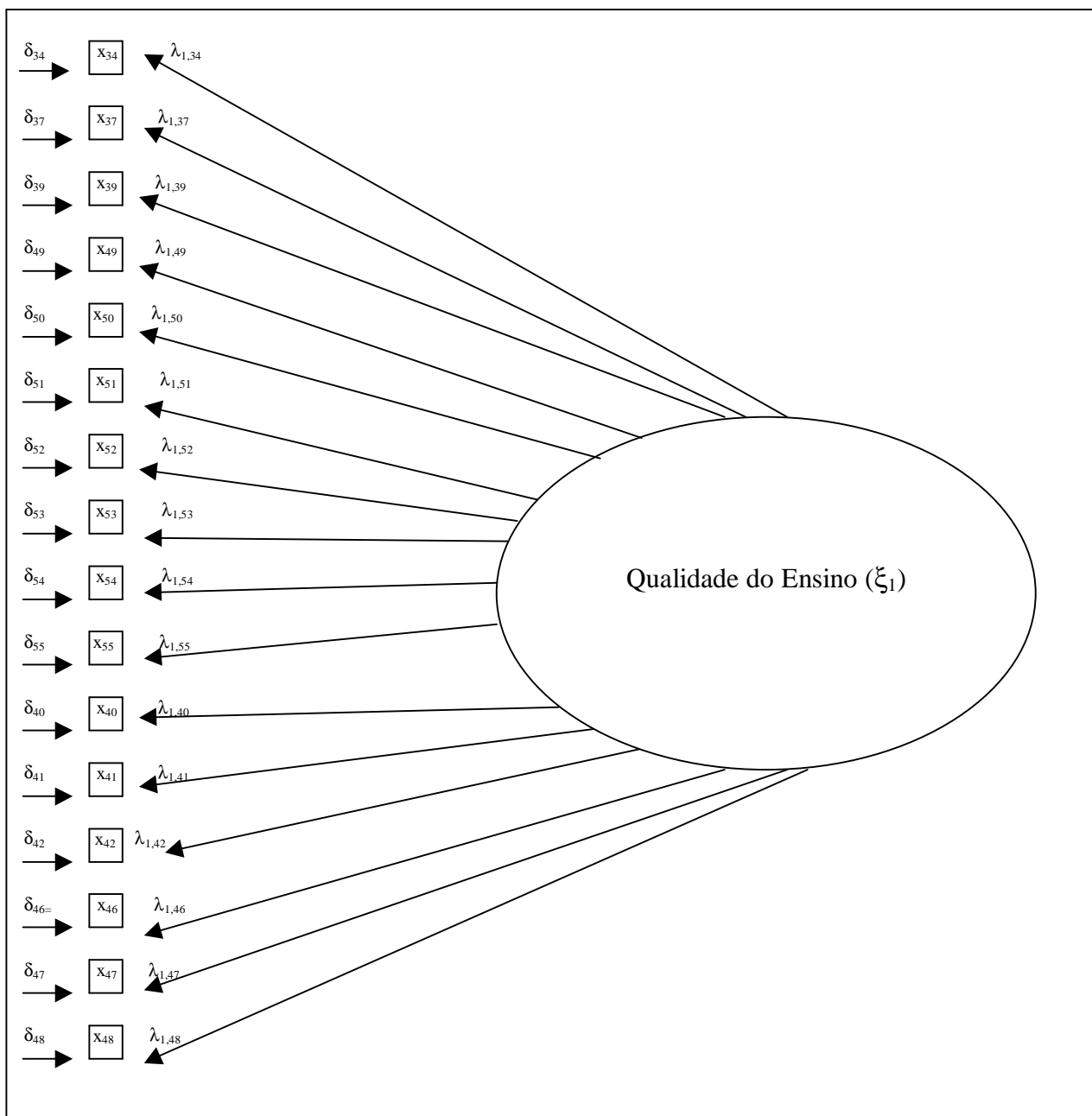


Diagrama 5: Diagrama de caminhos das características da qualidade docente - um fator latente.

As sub-hipóteses 8.1 e 8.2, enunciadas a seguir, são necessárias para a comprovação das propriedades fundamentais do instrumento de medida, ou seja, as propriedades de confiabilidade e de validade. A não rejeição dessas sub-hipóteses conduz à aceitação da hipótese 8, e evidencia-se o suporte empírico a favor das propriedades de confiabilidade e validade da dimensão operacionalizada no instrumento, ou seja, a validade de construto.

Sub-hipótese 8.1: as avaliações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre letivo de 2002, na UNISUL e estratificadas segundo o *campus*, constituem uma base de dados para o suporte empírico da unidimensionalidade do instrumento de avaliação.

Sub-hipótese 8.2: a dimensão, decorrente da análise confirmatória, no geral da UNISUL e estratificada por *campus* é consistente conforme os critérios psicométricos.

Hipótese 9: o modelo qualitativo teórico apresentado no diagrama 6 é composto por quatro fatores correlacionados, confiáveis e válidos da qualidade do ensino no âmbito da sala de aula.

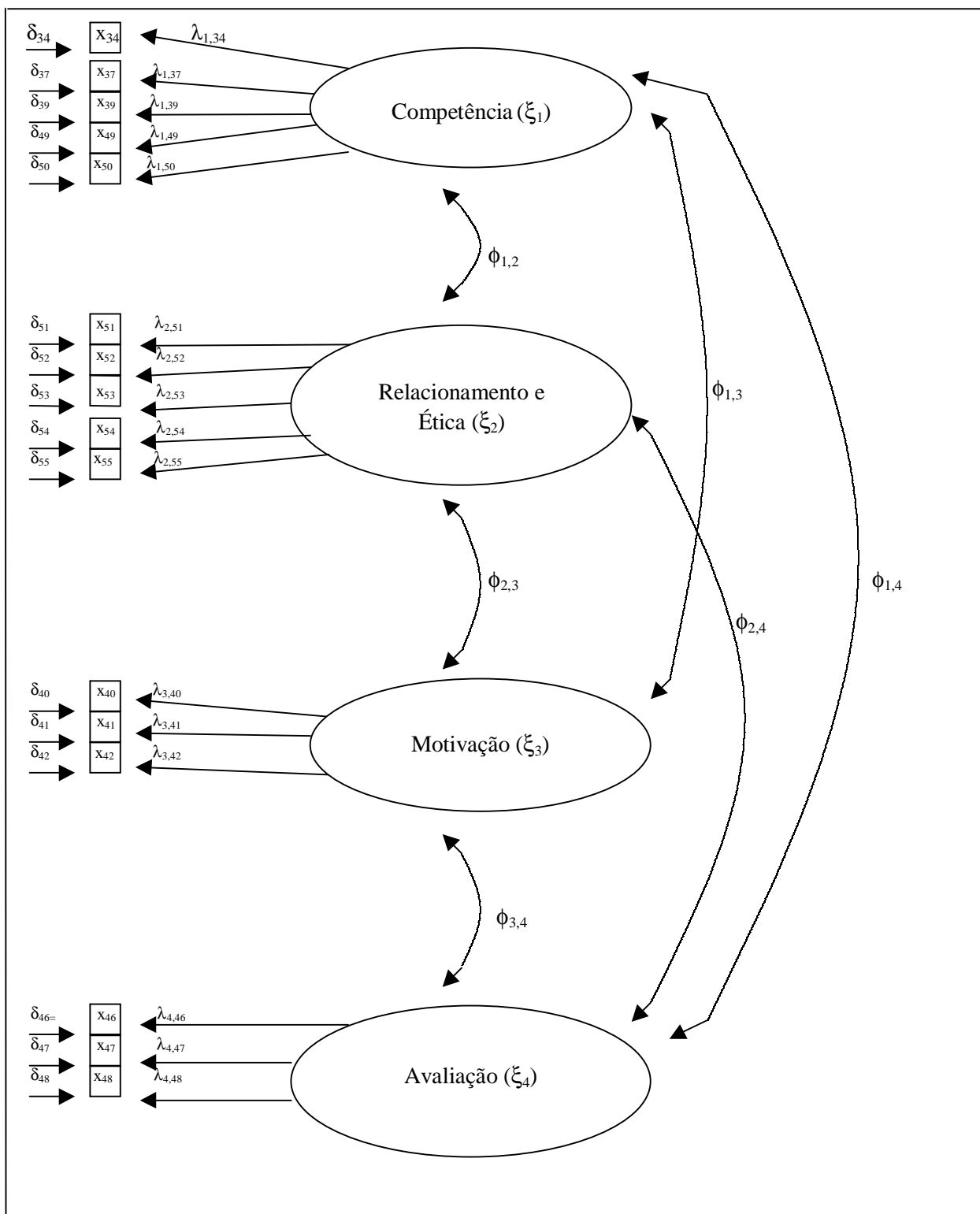


Diagrama 6: Diagrama de caminhos das características da qualidade docente com quatro fatores latentes correlacionados.

Na estimação do modelo com quatro fatores correlacionados, as seguintes suposições são necessárias: a) cada indicador apresenta carga fatorial significativa sobre um único fator; b) os quatro fatores são correlacionados entre si; e c) não existe correlação entre os resíduos resultantes da modelagem.

As sub-hipóteses 9.1 e 9.2 proferidas abaixo, são necessárias para a comprovação das propriedades estatísticas de confiabilidade e validade do instrumento em análise. A não rejeição dessas sub-hipóteses conduz à aceitação da hipótese 9, e evidencia-se o suporte empírico a favor das propriedades de confiabilidade e validade dos fatores latentes, operacionalizadas no instrumento; ou seja, a validade de construto.

Sub-hipótese 9.1: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre letivo de 2002, na UNISUL e estratificadas segundo o *campus*, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento em análise.

Sub-hipótese 9.2: os fatores latentes correlacionados, decorrentes das análises confirmatórias, no geral da UNISUL e estratificados por *campus* são consistentes conforme os critérios psicométricos.

8.7 Estimação do modelo com um fator latente e do modelo com quatro fatores latentes correlacionados

O uso da matriz de correlação, para estimar o modelo de medida com um fator latente (diagrama 5) e do modelo com quatro fatores latentes correlacionados (diagrama 6), possibilita comparações diretas dentro desses modelos.

Quanto ao tamanho das amostras utilizadas nessas modelagens, a tabela abaixo (28), apresenta o total de disciplinas avaliadas, bem como o tamanho das amostras utilizadas para a

estimação das correlações a serem usadas na estimação dos parâmetros desses modelos, para o geral UNISUL e estratificados segundo o *campus*.

Tabela 28: Tamanho das amostras utilizadas para comprovação dos modelos apresentados nos diagramas 5 e 6.

UNISUL	População	Tamanho da amostra (modelo com único fator latente)*	Tamanho da amostra (modelo com 4 fatores latente)
Tubarão	956	240	330
Araranguá	249	249	249
Pedra Branca	837	240	330
Geral	2042	240	330

* As matrizes de correlações utilizadas nas estimações dos modelos 10.4 e 10.5, estão apresentadas nos anexos X, XI, XII, XIII, XIV e XV.

Na determinação do tamanho das amostras foram levados em consideração a sensibilidade do método de estimação (máxima verossimilhança) para grandes amostras, o tamanho da população e o fato de que os dados empíricos, mesmo transformados, não apresentem coeficiente de normalidade multivariada de Nardia igual a zero. Segundo recomendações de Hair *et al.* (1999), a falta de normalidade multivariada dos dados empíricos eleva de 10 para 15 o número de avaliadores por parâmetro a ser estimado.

Após ter especificado que os modelos de medida com único fator latente e com quatro fatores latentes podem ser estimados, utilizando-se de mediadas confiáveis, é avaliada a capacidade dos modelos para estimar os parâmetros únicos dos modelos de medida, ou seja, avalia-se o seu nível de identificação.

A tabela abaixo 29 relata os números de parâmetros e de correlações a serem estimados, totais de correlações a serem utilizadas nessas modelagens e o número de graus de liberdade necessário para estimar os parâmetros dos modelos, corretamente.

Tabela 29: Graus de liberdade dos modelos de medida apresentados nos diagramas 5 e 6.

Modelos	Nº de parâmetros	Correlações entre fatores	Total de correlação	Graus de liberdade
Diagrama 5	16	0	120	104
Diagrama 6	16	6	120	98

Como os modelos de medidas citados na tabela acima apresentam um número positivo de graus de liberdade, é possível estimar os dois modelos de medida com indicadores de ajustes corretos e estimações únicas dos parâmetros não conhecidos, com graus de liberdade suficientes.

Utilizando-se da técnica de minimização de máxima verossimilhança, num primeiro momento, estima-se o modelo nulo; ou seja, o modelo em que todos os indicadores da escala estão associados a um único fator (diagrama 5). A seguir, estima-se o modelo de medida com quatro fatores correlacionados (diagrama 615). Ambos os modelos serão estimados para o geral UNISUL e estratificados por *campus*. Por conseguinte, oito modelos são estimados e avaliados. Os índices de adequação do ajuste do modelo nulo são confrontados com os índices de adequação do ajuste do modelo com quatro fatores correlacionados em todos os casos.

Segundo Bentler (1978, *apud* RAMOS, 1987) ao assegurar a validade de construto por meio de modelo causal, o grau de ajuste entre o modelo causal e os dados empíricos é o indicativo do grau de validade de construto. Assim, quando os índices de ajuste absoluto, incremental e de parcimônia apresentarem valores que satisfaçam, respectivamente, seus limites teóricos, maiores são as evidências do grau de validade dos construtos operacionalizados no instrumento de medida; ou seja, o instrumento mede corretamente os conceitos teóricos para os quais foram designados a medir.

8.8 Estimação do modelo com um fator latentes (modelo nulo) e do modelo com quatro fatores latente correlacionados – geral UNISUL

Como produto da estimação por máxima verossimilhança (ML), tem-se os índices de qualidade do ajuste apresentados na tabela 30 a seguir:

Tabela 30: 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

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

Quanto aos índices de ajuste absoluto, o modelo com quatro fatores correlacionados é o que apresenta melhores resultados em todos os tipos de índices. O valor do erro de aproximação quadrático médio (RMSE) 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 4, 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 significância estatística de 99,73%.

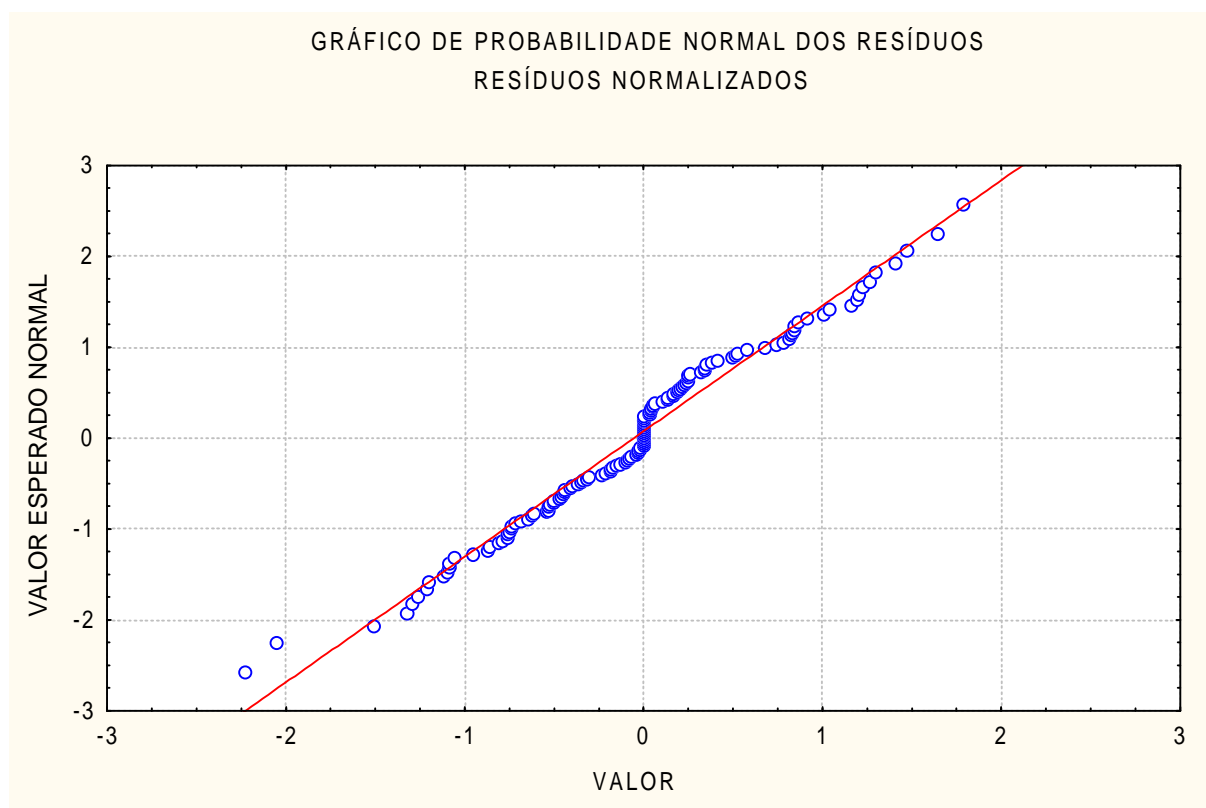


Gráfico 4: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados -geral UNISUL.

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, globalmente, prevê satisfatoriamente a matriz observada.

Em relação aos índices incrementais de ajuste, também o modelo com quatro fatores correlacionados é o que se obtêm melhores resultados. Os índices TLI e BBI, respectivamente, iguais a 0,933 e 0,938 são superiores ao do modelo nulo, ou seja, aquele que se supõe que todos os indicadores estão saturados em um único fator latente. 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 são os que possuem mais sentidos dentro de uma estratégia de modelização competitiva por oferecer medidas de ajuste do modelo em função dos números de coeficientes estimados. Os valores desses índices 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 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 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 31. A parte inferior da tabela apresenta a matriz de correlação entre os quatro fatores com os índices de confiabilidade composta na diagonal principal.

Quanto à significância dos indicadores, todos são positivamente e significativamente relacionados com seus fatores latentes específicos. Esse fato pode ser constatado dividindo cada λ estimado pelo seu respectivo erro padrão. Todos os *t student*, assim obtidos, 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.

Quanto à confiabilidade dos indicadores, todos apresentam erros de medida menores que 0,5. Todos os indicadores exibem confiabilidades bastante elevadas, indicando que eles explicam uma alta porcentagem da variância das variáveis latentes. A confiabilidade composta de cada fator latente é muito elevada, superando em todos os casos o limite mínimo de aceitação de 0,7. Desse modo, conclui-se 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 associado a esse fator latente é suficientemente explicada por esse fator, em todos os fatores latentes em análise.

Em síntese, há evidências de que o modelo de medida com quatro fatores correlacionados – geral UNISUL ajusta-se satisfatoriamente aos dados, possibilitando concluir que os indicadores considerados para os fatores latentes são adequados e confiáveis.

Tabela 31: Análise fatorial confirmatório. 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	ξ_2	ξ_3	ξ_4
ξ_1	0,954			
ξ_2	0,768	0,942		
ξ_3	0,689	0,637	0,872	
ξ_4	0,803	0,820	0,581	0,922

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade composta do fator.

* Todos os lâmbdas (λ) são estatisticamente significativo ao nível de 1%.

** Variância extraída da variável latente.

8.9 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – *campus* Tubarão

Como produto da estimação por máxima verossimilhança (ML) do modelo com um fator latente e do modelo com quatro fatores latentes correlacionados, tem-se para o *campus* de tubarão, os seguintes índices da qualidade de ajuste apresentados na tabela (32) abaixo.

Tabela 32: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores

Medidas de bondade do ajuste	Campus: Tubarão	
	Modelo com único fator latente(modelo nulo)	Modelo com 4 fatores correlacionados
Absoluto de ajuste		
χ^2	1181,78 (gl 104)	500,167 (gl 98)
RMSER	0,252	0,114 (0,105; 0,124)
RMSR	0,091	0,058
GFI	0,532	0,835
PNC	1077,78	402,167
PNCE	4,491	1,183
Incrementais de ajuste		
AGFI	0,338	0,771
TLI	0,700	0,910
BBI	0,723	0,910
Ajuste de parcimônia		
PNFI	0,627	0,743
PGFI	0,461	0,682
χ^2 (Normalizado)	11,363	5,104
AIC	5,212	1,751

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

Quanto aos índices de ajuste absoluto, o modelo com quatro fatores correlacionados é o que apresenta melhores resultados. O valor do erro de aproximação quadrático médio (RMSE) é de 0,114, podendo estender-se entre 0,105 e 0,124 com nível de confiança de 90%. A raiz quadrada média dos resíduos é de 0,058. O gráfico de probabilidade normal para os resíduos padronizados, apresentados no gráfico 5, exhibe todos os resíduos padronizados no intervalo de -3 a +3, o que indica que a matriz de correlação prevista, não se distanciam da matriz de correlação observada com confiabilidade estatística de 99,73%.

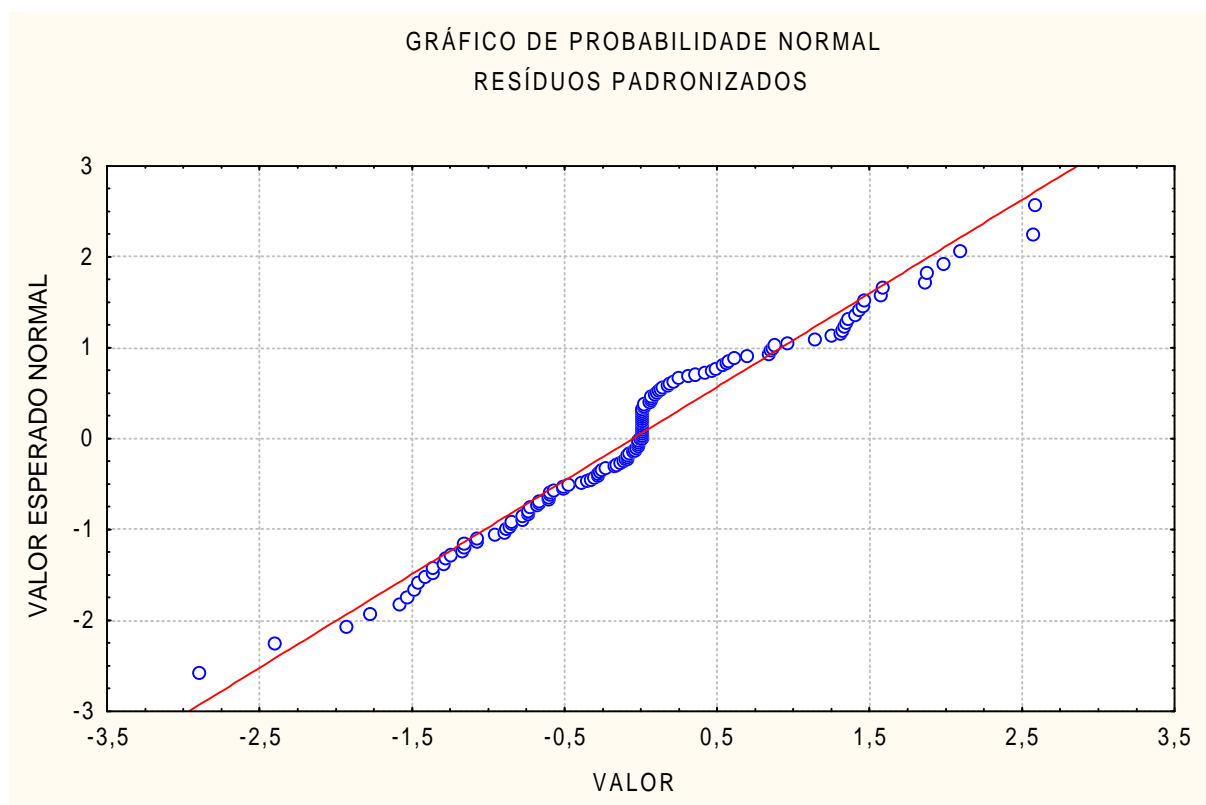


Gráfico 5: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados - campus Tubarão.

O índice GFI indica que 83,5% da variabilidade dos indicadores saturados pelos seus respectivos construtos são por ele explicado. Em relação aos índices PNC e PNCE, verifica-se que são maiores no modelo com quatro fatores correlacionados. Pode-se concluir,

então, que o modelo com quatro fatores correlacionados, globalmente, prevê satisfatoriamente a matriz observada.

O modelo com quatro fatores correlacionados, em relação aos índices incrementais de ajuste, é o que obtêm os melhores resultados. Os índices TLI e BBI, respectivamente, iguais a 0,910 e 0,910 são superiores ao do modelo nulo, ou seja, aquele que se supões que todos os indicadores estão saturados em um único fator latente. 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.

Quanto aos índices de parcimônia, estes confirmam o modelo com quatro fatores correlacionados como o mais adequado para reproduzir a matriz de correlação da amostra.

Aceito que os dados empíricos se ajustam de maneira satisfatória ao modelo proposto de quatro fatores latentes correlacionados, efetua-se a análise detalha de cada construto, observando a significância de cada indicador, a quantidade de erro de medida, a confiabilidade dos indicadores e dos construtos. Todos esses indicadores da qualidade do ajuste estão apresentados na tabela 33. A parte inferior da tabela apresenta a matriz de correlação entre os quatro fatores latentes com os índices de confiabilidade composta na diagonal principal.

Em relação à significância dos indicadores, pode-se afirmar que todos são positivamente e significativamente relacionados com seus fatores latentes respectivos. Esse fato pode ser constatado, dividindo cada λ , estimado pelo seu respectivo erro padrão. Todos os *t student*, assim obtidos, são estatisticamente significativos em níveis menores que 1%. Dessa forma, todos os indicadores estão significativamente relacionados com seus construtos específicos, comprovando a relação indicada entre os indicadores e os construtos.

No que diz respeito à confiabilidade dos indicadores, todos exibem erros de medida menores que 0,5, e confiabilidades superiores a 0,5. Essas altas confiabilidades indicam que os indicadores medidos explicam uma alta porcentagem da variância das variáveis latentes. As confiabilidades compostas dos quatro fatores latentes são elevadas, superando, em todos os casos, o limite de aceitação de 0,7. Desse modo, afirma-se, então, 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 desse fator latente é suficientemente explicada por esse fator.

Em síntese, há evidências de que o modelo de medida de indicadores exógenos, *campus* de Tubarão, ajusta-se de maneira satisfatória aos dados, possibilitando concluir que os indicadores indicados para os construtos são adequados e confiáveis.

Tabela 33: Análise fatorial confirmatório. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – Campus de Tubarão.

Fator	Lâmbda (λ^*)	Erro padrão	Theta delta (δ)	Confiabilidade do indicador
Competência (ξ_1)				
X ₃₄	0,889	0,013	0,210	0,790
X ₃₇	0,864	0,015	0,253	0,746
X ₃₉	0,900	0,012	0,190	0,810
X ₄₉	0,905	0,011	0,181	0,819
X ₅₀	0,965	0,006	0,068	0,931
	0,784**			
Relacionamento/ ética (ξ_2)				
X ₅₁	0,845	0,018	0,287	0,714
X ₅₂	0,869	0,016	0,244	0,755
X ₅₃	0,885	0,014	0,216	0,783
X ₅₈	0,793	0,022	0,371	0,629
X ₆₀	0,889	0,014	0,210	0,790
	0,685**			
Exigência (ξ_3)				
X ₄₀	0,831	0,023	0,310	0,691
X ₄₁	0,781	0,026	0,390	0,610
X ₄₂	0,881	0,020	0,224	0,776
	0,692**			
Avaliação (ξ_4)				
X ₄₆	0,870	0,016	0,243	0,757
X ₄₇	0,837	0,019	0,300	0,701
X ₄₈	0,974	0,010	0,104	0,949
	0,788**			
Valores phi				
	ξ_1	ξ_2	ξ_3	Ξ_4
ξ_1	0,958			
ξ_2	0,732	0,932		
ξ_3	0,633	0,582	0,871	
ξ_4	0,768	0,805	0,511	0,917

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade composta do fator.

*Todos os lâmbdas (λ) são estatisticamente significativo ao nível de 1%.

** Variância extraída do fator latente.

8.10 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latentes correlacionados – *campus* Araranguá

Apresenta-se, a seguir, como produto da estimação por máxima verossimilhança (ML), do modelo com único fator latente e do modelo com quatro fatores latentes correlacionados, para o *campus* de Araranguá, os seguintes índices da qualidade de ajuste, apresentados na tabela 34.

Tabela 34: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores.

Medidas de adequação do ajuste	Campus: Araranguá	
	Modelo com único fator latente(modelo nulo)	Modelo com 4 fatores correlacionados
Absoluto de ajuste		
χ^2	1307,01 (gl 104)	371,303 (gl 98)
RMSER	0,266	0,106 (0,094; 0,117)
RMSR	0,088	0,0431
GFI	0,507	0,843
PNC	1203,01	273,303
PNCE	4,831	1,098
Incrementais de ajuste		
AGFI	0,355	0,782
TLI	0,689	0,919
BBI	0,715	0,925
Ajuste de parcimônia		
PNFI	0,619	0,750
PGFI	0,439	0,688
χ^2 (Normalizado)	12,567	3,789
AIC	5,528	1,804

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

Em relação aos índices de ajuste absoluto, pode-se afirmar que o modelo com quatro fatores correlacionados é o que apresenta melhores resultados. Quanto ao valor do erro de aproximação quadrático médio (RMSE) é de 0,106 e pode se alongar entre 0,094 e 0,117 com nível de confiança de 90%. A raiz quadrada média dos resíduos é de 0,0431. O gráfico de probabilidade normal para os resíduos padronizados, visualizado no gráfico 6, exibe todos os resíduos padronizados no intervalo de -3 a +3, indicando indica que a matriz de correlação prevista, não se distancia da matriz de correlação observada com confiabilidade estatística de 99,73%.

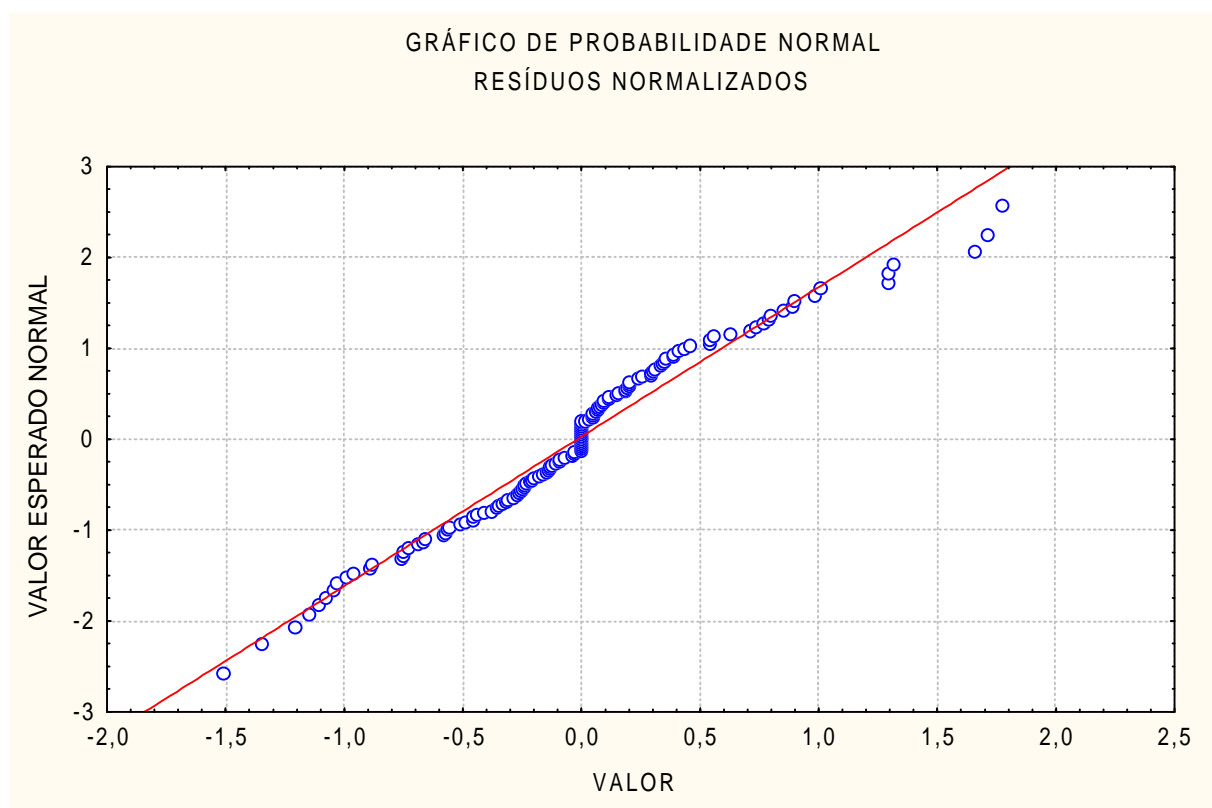


Gráfico 6: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados - campus Araranguá.

O índice GFI indica que 84,3% da variabilidade dos indicadores são explicados pelas quatro dimensões latentes. Quanto aos índices PNC e PNCE, esses também são maiores no modelo com quatro fatores correlacionados. Pode-se afirmar que o modelo com quatro fatores correlacionados, globalmente, pode prever de maneira satisfatória a matriz de correlação empírica.

Em se tratando do modelo com quatro fatores correlacionados, em relação aos índices incrementais de ajuste, é o que obtém os melhores resultados. Os índices TLI e BBI, respectivamente, iguais a 0,919 e 0,925 são superiores ao do modelo nulo. Em resumo, 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 são os que possuem mais sentidos dentro de uma estratégia de modelização competitiva, por oferecer medidas de ajuste do modelo em função do número de coeficientes estimados. Os valores desses índices confirmam o modelo com quatro fatores correlacionados como o mais adequado para reproduzir a matriz de correlação da amostra. Destaca-se, nesse bloco, o critério de informação de *Akaike*, adequado para comparar modelos com diferentes números de construtos. O valor próximo de zero indica que o modelo estimado é parcimonioso.

Depois de confirmado que os dados empíricos se ajustam de maneira satisfatória ao modelo proposto de quatro fatores latentes correlacionados, parte-se para a análise detalhada de cada construto. Para isso, deve-se observar a significância de cada indicador, a quantidade de erro de medida, a confiabilidade dos indicadores e dos construtos.

Na tabela 35, pode-se visualizar o coeficiente padronizado, o erro padrão de estimativa, o erro de medida e a confiabilidade para cada um dos indicadores, bem como a

variância extraída de cada fator. A parte inferior da tabela apresenta a matriz de correlação entre os quatro fatores com os índices de confiabilidade composta na diagonal principal.

No que diz respeito à significância dos indicadores, todos estão positivamente e significativamente relacionados com seus fatores latentes respectivos. Pode-se constatar esse fato, dividindo cada λ 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 construtos específicos, comprovando a relação proposta entre os indicadores e os construtos.

Todos os indicadores exibem erros de medida menores que 0,5, e confiabilidades superiores a 0,5. Dessa forma, pode-se afirmar que essas altas confiabilidades indicam que os indicadores medidos explicam uma alta porcentagem da variância das variáveis latentes. As confiabilidades compostas dos quatro fatores latentes são elevadas, superando em todos os casos o limite de aceitação de 0,7. Assim, conclui-se 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 desse fator latente é suficientemente explicada por ele.

Há evidências, dessa forma, de que o modelo de medida de indicadores exógenos, *campus* de Araranguá, ajusta-se de maneira satisfatória aos dados, possibilitando concluir que os indicadores considerados para o construto são adequados e confiáveis.

Tabela 35: Análise fatorial confirmatório. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do modelo com quatro fatores correlacionados – Campus de Araranguá.

Fator	Lâmbda (λ^*)	Erro padrão	Theta delta (δ)	Confiabilidade do indicador
Competência (ξ_1)				
X ₃₄	0,868	0,017	0,246	0,753
X ₃₇	0,903	0,013	0,185	0,815
X ₃₉	0,906	0,013	0,180	0,821
X ₄₉	0,896	0,014	0,196	0,803
X ₅₀	0,947	0,008	0,102	0,897
	0,818**			
Relacionamento/ ética (ξ_2)				
X ₅₁	0,890	0,015	0,208	0,792
X ₅₂	0,880	0,016	0,226	0,774
X ₅₃	0,883	0,016	0,220	0,780
X ₅₈	0,871	0,017	0,241	0,759
X ₆₀	0,933	0,010	0,130	0,870
	0,758**			
Exigência (ξ_3)				
X ₄₀	0,909	0,020	0,174	0,826
X ₄₁	0,753	0,032	0,433	0,567
X ₄₂	0,840	0,025	0,294	0,706
	0,700**			
Avaliação (ξ_4)				
X ₄₆	0,898	0,015	0,194	0,806
X ₄₇	0,887	0,016	0,213	0,787
X ₄₈	0,929	0,012	0,138	0,863
	0,545**			
Valores phi	ξ_1	ξ_2	ξ_3	ξ_4
ξ_1	0,957			
ξ_2	0,751	0,951		
ξ_3	0,679	0,668	0,847	
ξ_4	0,796	0,842	0,610	0,931

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade composta do fator.

*Todos os lâmbdas (λ) são estatisticamente significativo ao nível de 1%.

** Variância extraída da variável latente.

8.11 Estimação do modelo com um fator latente (modelo nulo) e do modelo com quatro fatores latente correlacionados – *campus* Pedra Branca

Com relação ao *campus* de Pedra Branca, como produto da estimação por máxima verossimilhança (ML), do modelo com um fator latente e do modelo com quatro fatores

latentes correlacionados, os seguintes índices da qualidade de ajuste segue apresentado na tabela 36.

Tabela 36: Índices de adequação do ajuste do modelo nulo e do modelo com quatro fatores correlacionados. Estimação por máxima verossimilhança – Campus: Pedra Branca.

Medidas de adequação do ajuste	Campus: Pedra Branca	
	Modelo com único fator latente(modelo nulo)	Modelo com 4 fatores correlacionados
Absoluto de ajuste		
χ^2	1098,44 (gl 104)	365,66 (gl 98)
RMSER	0,243	0,090 (0,081; 0,100)
RMSR	0,082	0,039
GFI	0,548	0,880
PNC	994,440	258,66
PNCE	4,144	0,784
Incrementais de ajuste		
AGFI	0,409	0,833
TLI	0,739	0,940
BBI	0,756	0,945
Ajuste de parcimônia		
PNFI	0,656	0,767
PGFI	0,475	0,719
χ^2 (Normalizado)	10,562	3,731
AIC	4,864	1,303

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

O modelo com quatro fatores correlacionados, quanto aos índices de ajuste absoluto, é o que apresenta melhores resultados. O valor do erro de aproximação quadrático médio (RMSER) é de 0,090, podendo se alongar entre 0,081 e 0,100, com nível de confiança de 90%. A raiz quadrada média dos resíduos é de 0,039. O gráfico de probabilidade normal para os resíduos padronizados, que pode ser visualizado no gráfico 7, exibe todos os resíduos padronizados no intervalo de -3 a +3. Indica dessa forma, que a matriz de correlação prevista não se distancia da matriz de correlação observada com confiabilidade estatística de 99,73%.

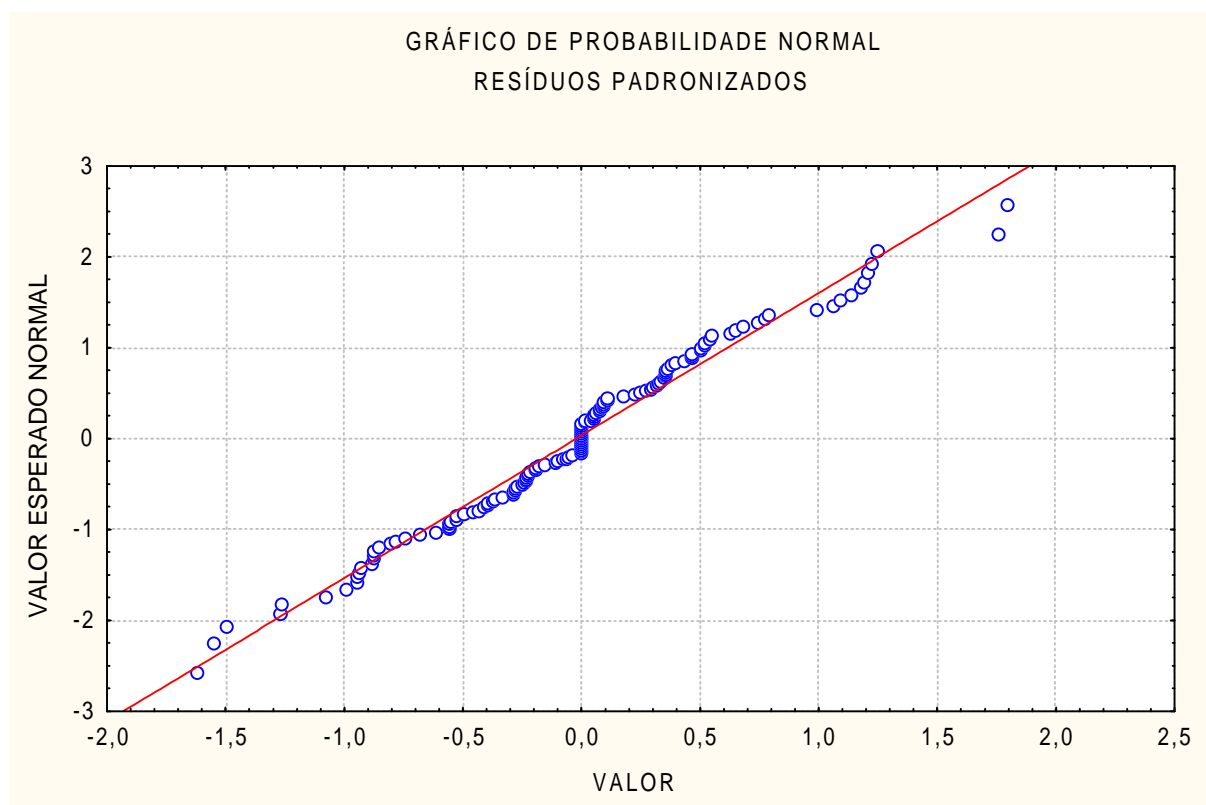


Gráfico 7: Gráfico de probabilidade normal para os resíduos do modelo com quatro fatores correlacionados -campus Pedra Branca.

O índice GFI indica que 88,0% da variabilidade dos indicadores são explicados pelas quatro dimensões latentes. Quanto aos índices PNC e PNCE, esses também são maiores no modelo com quatro fatores correlacionados. Diante do exposto, conclui-se que o modelo com quatro fatores correlacionados, globalmente, prever de maneira suficiente a matriz observada.

Pode-se afirmar que em relação aos índices incrementais de ajuste, o modelo com quatro fatores correlacionados é o que obtém os melhores resultados. Os índices TLI e BBI, respectivamente, iguais a 0,940 e 0,945 são superiores ao do modelo nulo. Em resumo, 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, conforme já mencionado anteriormente, são os que possuem mais sentidos dentro de uma estratégia de modelização competitiva em virtude de oferecer medidas de ajuste do modelo em função do número de coeficientes estimados. Desta forma, pelos valores desses índices, pode-se afirmar que o modelo com quatro fatores correlacionados é o mais adequado para reproduzir a matriz de correlação da amostra. Convém lembrar que o critério de informação de *Akaike* é adequado para comparar modelos com diferentes números de construtos. Valor próximo de zero indica que o modelo estimado é parcimonioso.

Após verificar-se que os dados empíricos se ajustam ao modelo proposto de quatro fatores latentes correlacionados, passa-se então para a análise detalhada de cada construto. Deve-se observar a significância de cada indicador, a quantidade de erro de medida, a confiabilidade dos indicadores e dos construtos.

Para melhor visualizar, apresenta-se, na tabela 37, o coeficiente padronizado, o erro padrão de estimativa, o erro de medida e a confiabilidade para cada um dos indicadores e

também a variância extraída de cada fator. Na parte inferior da tabela, apresenta-se a matriz de correlação entre os quatro fatores com os índices de confiabilidade composta na diagonal principal.

No que tange à significância dos indicadores, pode-se afirmar que todos são positivamente e significativamente relacionados com seus fatores latentes respectivos. Conforme já citado anteriormente, este fato pode ser constatado, dividindo cada λ estimado pelo seu respectivo erro padrão. Todos os *t student*, assim obtidos, são estatisticamente significativos em níveis menores que 1%. Assim, todos os indicadores estão significativamente relacionados com seus construtos específicos, comprovando a relação proposta entre os indicadores e os construtos.

Com relação à confiabilidade dos indicadores, conforme pode ser visualizado na tabela 35, todos exibem erros de medida menos que 0,5, e confiabilidades superiores a 0,5. Por apresentar essas altas confiabilidades os indicadores medidos explicam uma alta porcentagem da variância das variáveis latentes. As confiabilidades compostas dos quatro fatores latentes são elevadas, superando em todos os casos o limite 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 desse fator latente é suficientemente explicada por ele.

Como se pôde observar com os modelos aplicados no *campus* de Tubarão e no *campus* de Araranguá, há evidências de que o modelo de medida de indicadores exógenos, também no *campus* de Pedra Branca, ajusta-se de maneira satisfatória aos dados, possibilitando concluir que os que os indicadores considerados para o construto, são adequados e confiáveis.

Tabela 37: Análise fatorial confirmatória. Estimadores padronizados, erros padrões, erros de medida, variância extraída, correlação entre fatores e confiabilidade composta do fator. Estimação por máxima verossimilhança – Modelo com quatro fatores correlacionados - *Campus Pedra Banca*.

Fator	Lâmbda (λ^*)	Erro padrão	Theta delta (δ)	Confiabilidade do indicador
Competência (ξ_1)				
X ₃₄	0,897	0,012	0,196	0,805
X ₃₇	0,863	0,015	0,254	0,745
X ₃₉	0,901	0,011	0,189	0,812
X ₄₉	0,891	0,012	0,206	0,794
X ₅₀	0,958	0,006	0,083	0,920
	0,779**			
Relacionamento/ ética (ξ_2)				
X ₅₁	0,897	0,012	0,196	0,805
X ₅₂	0,860	0,016	0,260	0,740
X ₅₃	0,906	0,011	0,179	0,821
X ₅₈	0,865	0,015	0,252	0,748
X ₆₀	0,915	0,011	0,163	0,837
	0,790**			
Exigência (ξ_3)				
X ₄₀	0,878	0,018	0,229	0,771
X ₄₁	0,811	0,023	0,343	0,658
X ₄₂	0,861	0,019	0,259	0,741
	0,723**			
Avaliação (ξ_4)				
X ₄₆	0,900	0,013	0,190	0,810
X ₄₇	0,869	0,016	0,245	0,755
X ₄₈	0,910	0,012	0,172	0,828
	0,798**			
Valores phi	ξ_1	ξ_2	ξ_3	ξ_4
ξ_1	0,956			
ξ_2	0,777	0,949		
ξ_3	0,730	0,644	0,887	
ξ_4	0,830	0,813	0,606	0,922

Elemento em negrito da diagonal principal da matriz de correlação entre os fatores corresponde à confiabilidade composta do fator.

*Todos os lâmbdas (λ) são estatisticamente significativo ao nível de 1%.

** Variância extraída da variável latente.

8.12 Comparação dos índices de adequação do ajuste do modelo: geral UNISUL e de outros pesquisadores.

Ao se comparar os índices de qualidade do ajuste dos instrumentos de avaliação da docência universitária de Ramos (1987) de Capelleras e Veciana (2000) com os obtidos pelo modelo com quatro fatores latentes – geral UNISUL (admitindo que o modelo de medida com quatro fatores latentes correlacionados, generalize-se nos campi), relatam-se os seguintes índices apresentados na tabela 38.

Tabela 38: Comparação dos índices de adequação do ajuste.

Instrumentos	Nos de fatores correlacionados	Índices de qualidade do ajuste			
		GFI	AGFI	RMSE	RMSR
Ramos (1997)	6	0,928	0,900	-	0,037
Capelleras e Veciana (2000)	5	0,902	0,875	0,069	-
UNISUL	4	0,873	0,824	0,102	0,044

Inicialmente estimou-se o modelo de medida com oito fatores correlacionados, obtendo-se qui-quadrado de 3474,20 (tabela 17). A análise detalhada do modelo conduziu para a sua reformulação. Os estudos exploratórios evidenciaram, após a remoção de alguns indicadores, que traziam mais ruídos do que informações, que a estrutura da escala pode ser explicada com poucas perdas de informação, por quatro fatores latentes correlacionados. Essa hipótese foi submetida à prova através dos modelos causais e resultou no qui-quadrado de 375,235 (tabela 30). A proporção no aumento do ajuste, quando se compara o qui-quadrado do modelo com oito fatores correlacionados com o qui-quadrado do modelo com quatro fatores correlacionados, geral UNISUL, é aproximadamente 0,9. Bisquerra (1989, *apud*

GARCIA; MARTINEZ, 2000) propõe o valor 0,9 como o limite para interromper o processo de reformulação do modelo.

Os índices de adequação do ajuste absoluto e incremental reportados por Ramos (1997) e por Capelleras e Veciana (2000), reportados na tabela 38, indicam que seus modelos são levemente superiores ao modelo com quatro fatores latente correlacionados – geral UNISUL. Mas, considerando que o modelo com quatro fatores latentes (geral da UNISUL) não alcance melhores índices de ajuste e que validade é uma questão de grau, pode-se concluir que este modelo apresente pelo menos um ajuste aceitável e acolher que os discentes distinguem essas quatro dimensões da escala como relevantes para a efetividade do ensino, no âmbito da sala de aula. Essa conclusão confirma o modelo qualitativo teórico apresentado no quadro 15 e conseqüentemente todas as suas sub-hipóteses a ela relacionada, ou sejam:

Sub-hipótese 9.1: as valorações estudantis, atribuídas para o total de disciplinas cursadas no 2º semestre letivo de 2002, na UNISUL e estratificadas segundo o *campus*, constituem uma base de dados para o suporte empírico da multidimensionalidade do instrumento em análise; e

Sub-hipótese 9.2: os fatores latentes correlacionados, decorrentes das análises confirmatórias, no geral da UNISUL e estratificados por *campus* são consistentes conforme os critérios psicométricos.

8.13 Validez de critério da escala

A validade de critério refere-se ao grau de eficácia com que se pode prever um indicador a partir da medição efetuada. O instrumento de avaliação da UNISUL possui o indicador X_{61} , ou seja “indicaria este professor a outras turmas” que conforme figura 10.2,

apresenta elevada variabilidade, decorrente da existência de diferenças de opiniões entre os alunos de distintas disciplinas.

Segundo os discentes, os fatores competência, relacionamento e ética, motivação dos alunos e avaliação são os mais importantes, manifestando-se na variável critério “Indicaria o professor dessa disciplina a outras turmas” durante o processo de avaliação pelos discentes. Sendo esta característica verdadeira, o efeito desses fatores latentes sobre esta variável critério é positivo e significativo. A tabela a seguir apresenta os resultados da análise de regressão, tomando os escores fatoriais dos quatro fatores latentes obtidos pela análise fatorial como variáveis independentes e a variável critério “indicaria o professor da disciplina a outras turmas” como variável dependente.

Tabela 39: Análise de regressão.

Fatores	Indicaria esse professor a outras turmas.			
	Coefficiente Beta	Erro padrão	<i>t</i> (student)	p- valor
Competência	0,6940	0,008823	78,6374	0,0000
Relacionamento e ética	0,5225	0,008825	59,2039	0,0000
Motivação	0,1536	0,008825	17,400	0,0000
Avaliação	0,2434	0,008825	27,9203	0,0000
F(4, 2069) = 2692,8 ; p-valor =0,000			R ² (ajustado) = 83,86%	

Os valores de *t student* de todas as variáveis independentes e os níveis de significancia obtidos, assim como os coeficientes padronizados da regressão positivos, indicam que o modelo de regressão é estatisticamente significativo e que as quatro dimensões da escala influem de forma positiva e significativa na valoração da indicação desse professor a outras turmas. Portanto, confirma-se a hipótese quatro sobre o efeito desses fatores e evidenciada a validade de critério da escala.

Quanto à comprovação dos efeitos das dimensões validadas da escala sobre as variáveis critérios “considera que o seu curso tem qualidade” , “o seu curso está exigindo o suficiente para sua formação profissional”, não foi possível a sua total comprovação, uma vez que, nessa etapa, a unidade de análise é o curso.

No segundo semestre letivo de 2002 foram avaliados as disciplinas de 82 cursos de graduação. Ao considerar as médias dos cursos como unidades de análise, os dados são insuficientes, tanto para as análises exploratórias como confirmatórias. Assim, não foi possível comprovar a generalização das dimensões da escala nas médias dos indicadores dos cursos de graduação. No entanto, as correlações entre essas variáveis critérios e os indicadores das dimensões validadas da escala foram calculadas e testadas a sua significância e seguem apresentadas na tabela 40 abaixo.

Tabela 40: Correlações entre as variáveis critérios e os indicadores das dimensões validadas da escala.

Fator	X ₁	X ₄	X ₅	X ₆
Competência (ξ_1)				
X ₃₄	0,34*	0,35*	0,48*	0,50*
X ₃₇	0,46*	0,51*	0,47*	0,46*
X ₃₉	-0,09	-0,02	-0,02	0,01
X ₄₉	0,20	0,31*	0,25*	0,32*
X ₅₀	0,42*	0,50*	0,41*	0,36*
Relacionamento/ ética (ξ_2)				
X ₅₁	0,37*	0,39*	0,30*	0,30*
X ₅₂	0,06	0,17	0,03	0,12
X ₅₃	0,08	0,13	0,04	0,14
X ₅₈	0,20	0,26*	0,31*	0,38*
X ₆₀	0,28*	0,29*	0,18	0,21
Exigência (ξ_3)				
X ₄₀	0,19	0,27*	0,18	0,20
X ₄₁	0,28*	0,37*	0,17	0,26*
X ₄₂	0,20	0,25*	0,14	0,22*
Avaliação (ξ_4)				
X ₄₆	0,13	0,19	-0,00	0,06
X ₄₇	0,12	0,21	0,15	0,22*
X ₄₈	0,13	0,25*	0,13	0,30*

X₁ : Considera que o seu curso tem qualidade; X₄: O seu curso está exigindo o suficiente para sua formação profissional; X₅ : Considera a UNISUL uma universidade de Qualidade; X₆ : Você tem orgulho de estudar na UNISUL. * estatisticamente significativa ao nível de 5%.

Para Cashin (1988,1995), ao se analisar correlações de validade, correlações acima de 0,7 não são freqüentes em medidas educacionais, especialmente se for estudado fenômenos complexos, tais como o aprendizado do aluno. Esse autor sugere que os coeficientes de validade para os níveis de satisfação dos alunos entre 0,0 e 0,29, até quando estatisticamente significativos como inexistentes, coeficientes de validade entre 0,30 e 0,49 como sendo razoáveis e correlações entre 0,5 e 0,7 como boas.

Usando a escala de classificação acima e os resultados reportados na tabela 40, é possível concluir que os níveis de satisfação dos alunos sobre os diversos indicadores do ensino eficaz são válidos. Isso se deve ao fato de que todos os indicadores descritos no cabeçalho da tabela estarem razoavelmente correlacionados com os indicadores da dimensão “competência”, exceto com o indicador X₃₉. Destaca-se o indicador X₄ “o seu curso está

exigindo o suficiente para sua formação profissional”, que se correlaciona positivamente com a maioria dos indicadores da escala. Do exposto, conclui-se que as hipóteses (3) e (5) foram parcialmente comprovadas.

A comprovação da hipótese (6), que consiste na comprovação de que as dimensões da escala possuem um efeito positivo e significativo na valoração global da qualidade da disciplina percebida pelos alunos, está plenamente comprovada pelas relações causais exibida no diagrama 7.

Todas as relações causais, apresentadas no digrama 7, são positivas, significativas e de diferentes intensidades. O indicador X_{50} , ou seja “conduz o ensino da disciplina com qualidade”, é o indicador mais influenciado pelo fator “competência”, fato este comprovado pela elevada saturação desse indicador no fator. O indicador X_{60} , ou seja, “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} , ou seja, “estimula a utilização de bibliografia atualizada” é o indicador mais fortemente saturado no fator “motivação”, enquanto o indicador X_{48} , ou seja “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 medidas, logo comprovam as altas confiabilidades evidenciadas. As correlações entre os fatores latentes confirmam as obtidas na análise fatorial exploratória. Todas as correlações entre os fatores são positivas e significativas e evidenciam que o sucesso do professor depende de suas habilidades nesses fatores.

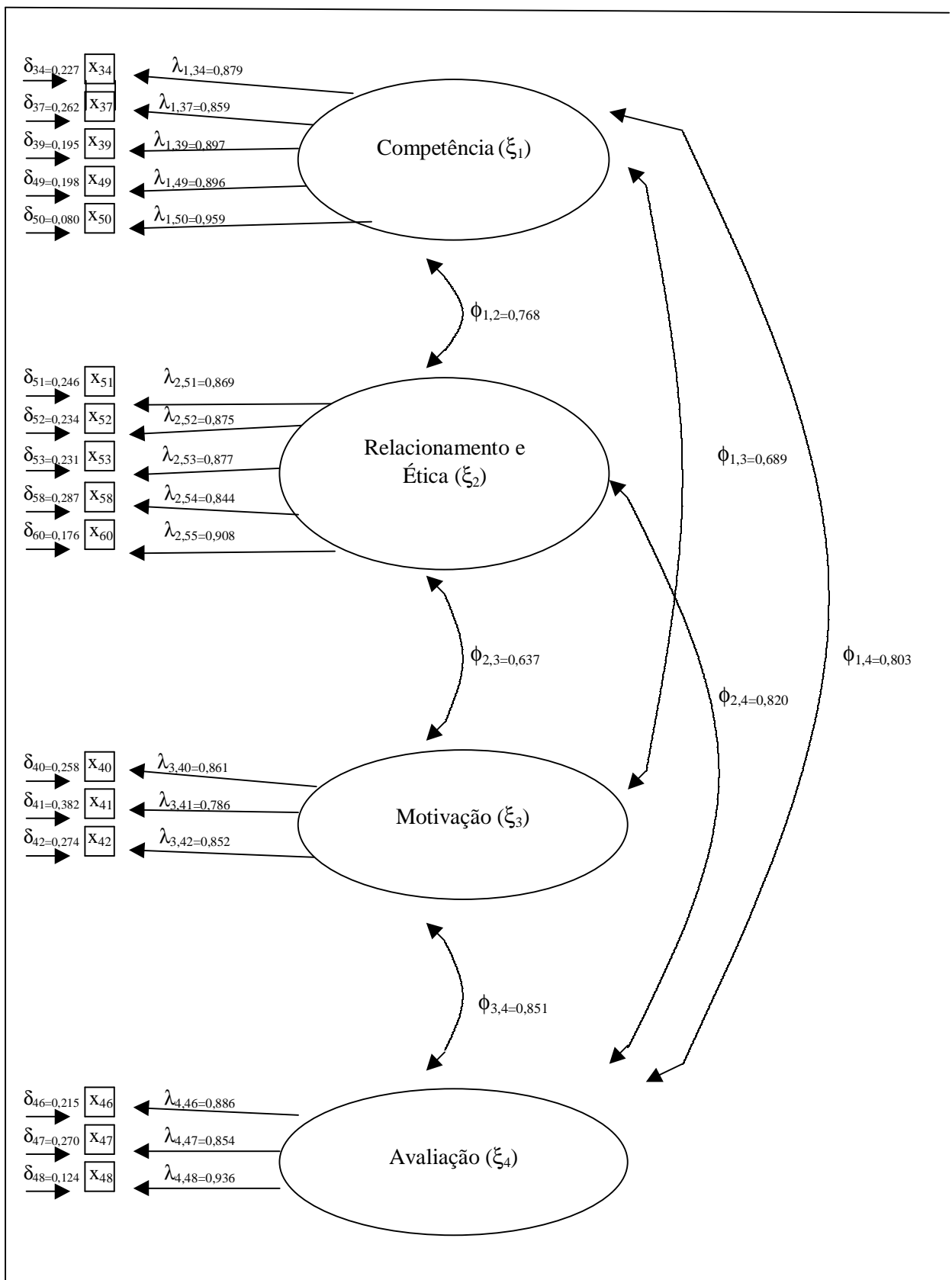


Diagrama 7: Diagrama de caminhos das características da qualidade docente – Geral UNISUL

8.14 Validade da escala para diagnosticar tipologias de docentes

Como último objetivo a ser alcançado neste estudo de caso é a capacidade do instrumento para diagnosticar tipologias de qualidade docente em função da satisfação que revelam os estudantes de graduação da UNISUL, e, desse modo, acumular mais evidências de validade para o instrumento em análise. Esse requisito é de fundamental importância para que possibilite um adequado uso das pontuações provenientes dos processos avaliativos. Da população de disciplinas avaliadas na UNISUL, no segundo semestre letivo de 2002, por limitação do software, foi obtida aleatoriamente uma amostra de 1750 disciplinas e realizada a análise de agrupamentos, utilizando a técnica não hierárquica K-médias. A solução de quatro agrupamentos possibilitou a discriminação apresentado no gráfico 8. Cabe salientar que o nível de satisfação dos discentes decresce em função do grupo de pertinência.

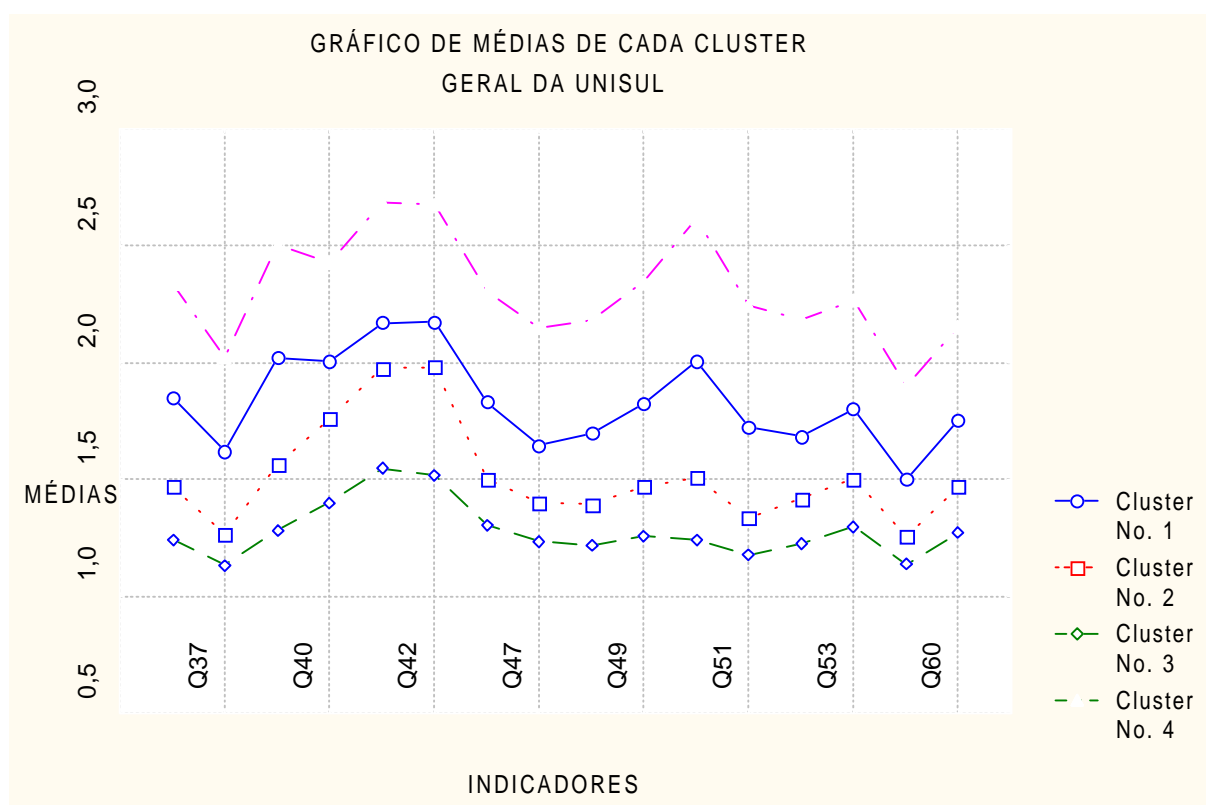


Gráfico 8: Médias dos grupos – geral UNISUL.

O agrupamento três é constituído pelas médias de 587 disciplinas que receberam as melhores avaliações nos dezesseis indicadores da escala. Neste agrupamento, e nos demais, os indicadores de números 40, 41 e 42 foram os que apresentaram piores desempenhos, enquanto os indicadores de números 37 e 54 foram as que receberam os melhores níveis de satisfação. O agrupamento dois é constituído por 657 disciplinas, que nos dezesseis indicadores foram avaliadas em torno de “praticamente sim”. O agrupamento um é composto 400 disciplinas que foram avaliadas em torno da categoria de resposta “praticamente não”, enquanto o agrupamento quatro é composto por 106 disciplinas que foram avaliadas em torno da categoria de resposta “não”.

Quanto à comprovação estatística das diferenças de médias, a tabela 41 abaixo relata que o teste de análise de variância comprova que o nível de significação é muito baixo em todos os indicadores; portanto, pode-se comprovar que as médias obtidas por estes indicadores diferem consideravelmente entre os grupos.

Tabela 41: Valores médios por cluster e análise de variância.

Indicadores	Cluster 3 (587 casos)	Cluster 2 (657 casos)	Cluster 1 (400 casos)	Cluster 4 (106 casos)	Estatística F (p _{valor})
X ₃₄	1,2442	1,4694	1,8519	2,3406	937,38 (0,00)
X ₃₇	1,1347	1,2666	1,6220	2,0237	808,08 (0,00)
X ₃₉	1,2822	1,5625	2,0190	2,5063	1472,69 (0,00)
X ₄₀	1,4033	1,7611	2,0056	2,4283	625,88 (0,00)
X ₄₁	1,5467	1,9780	2,1696	2,6861	442,14 (0,00)
X ₄₂	1,5177	1,9804	2,1758	2,6785	479,85 (0,00)
X ₄₆	1,3064	1,5010	1,8292	2,3043	588,66 (0,00)
X ₄₇	1,2360	1,4000	1,6474	2,1479	728,19 (0,00)
X ₄₈	1,2162	1,3947	1,6977	2,1812	819,45 (0,00)
X ₄₉	1,2577	1,4695	1,8250	2,3494	1127,25 (0,00)
X ₅₀	1,2422	1,5078	2,0067	2,6163	1864,42 (0,00)
X ₅₁	1,1763	1,3365	1,7246	2,2438	685,95 (0,00)
X ₅₂	1,2253	1,4189	1,6872	2,1843	979,36 (0,00)
X ₅₃	1,2982	1,5053	1,8014	2,2680	821,70 (0,00)
X ₅₈	1,1361	1,2562	1,5002	1,9008	433,06 (0,00)
X ₆₀	1,2728	1,4705	1,7565	2,1572	848,23 (0,00)

O resultados apresentados no gráfico 8 e na tabela 41 evidenciam que o questionário “Avaliação pelos Alunos do Desempenho da Disciplina” da UNISUL, reformulado para dezesseis indicadores, diferencia tipos de professores em função da satisfação que mostram os discentes. Esses estudos exploratórios acumulam mais evidências a favor do grau de validade da escala modificada.

A análise de agrupamento foi realizada com os dados transformados e não transformados com uma amostra de disciplinas obtidas aleatoriamente do total de disciplinas avaliadas no global da UNISUL, estratificadas por *campus* e por dimensões. Todas as análises são muito similares com as apresentadas no gráfico 8 e tabela 41.

9 CONCLUSÕES

Este trabalho teve como objetivo a apresentação de uma metodologia para testar as propriedades de confiabilidade e de validade da avaliação docente pelo discente, fundamentada principalmente em Marsh (1987), Ramos (1996) e Hair *et al.* (1999). A metodologia proposta está descrita nos capítulos 4, 5 e 6, que tratam da avaliação das propriedades do instrumento de avaliação mediante modelos de análise exploratória e confirmatória.

A metodologia foi utilizada visando obter as estimativas dos graus de confiabilidade e de validade do instrumento de avaliação utilizada pela equipe de Avaliação Institucional e Clima Organizacional da UNISUL. Para a comprovação dessas propriedades estatísticas do instrumento foi utilizada a base de dados proveniente do processo de avaliação de disciplinas ocorrido no segundo semestre letivo de 2002, totalizando 2511 disciplinas, avaliadas por 10802 (66%) alunos da UNISUL.

Foi proposta, nos capítulos 4 e 6 a abordagem de validação de construto, para confirmar as dimensões ou fatores latentes subjacentes ao instrumento de avaliação, uma vez que o questionário de avaliação da docência pelos alunos, utilizado pela instituição de ensino, engloba um conjunto de características que, *a priori*, definem um bom professor. Características estas que não são redutíveis umas as outras, portanto, torna o questionário multidimensional.

Assim sendo, o tipo de validade a ser evidenciado é o de validade de construto, ou seja, analisar através dos dados procedentes dos alunos que esses reconhecem as dimensões do instrumento de avaliação, isto é, confirmar que o instrumento mede algo concreto. Por outro lado, este tipo de validade engloba a validade convergente, preditiva, conteúdo e

discriminante. Isso quer dizer que, atribuído ao instrumento determinado grau de validade de construto, estar-se-á validando os demais tipos de validez.

No capítulo 5 é apresentada a técnica analítica de análise fatorial exploratória, uma vez que ajuda o pesquisador ou especialista em medição educacional na construção das dimensões do instrumento de avaliação. Este importante recurso de análises de dados, além de reduzir a dimensionalidade da escala, também informa sobre a validade e confiabilidade dos indicadores da escala. Sendo esta técnica, simplesmente exploratória, é necessário que o conjunto definido de relações seja modelado por modelo de análise fatorial confirmatória, caso particular dos modelos de equações estruturais. Os procedimentos analíticos necessários para validar os construtos operacionalizados na escala através da técnica de análise fatorial confirmatória são apresentados no capítulo 6.

Todos esses procedimentos foram utilizados para comprovar as propriedades estatísticas de confiabilidade e de validade do instrumento de avaliação da UNISUL. O referido instrumento foi designado, *a priori*, para medir oito fatores latentes ou dimensões da qualidade do ensino no domínio da sala de aula. Essa constitui a primeira hipótese que foi testada por meio dos procedimentos de análise fatorial confirmatória.

Com relação à hipótese de que as oito dimensões da escala apresentam grande capacidade de validade de discriminação, foi estimado o modelo com oito dimensões latentes não correlacionadas. Os indicadores de bondade do ajuste absoluto, incremental e de parcimônia foram todos considerados muito baixo sob o ponto de vista estatístico.

Admitindo que a capacidade de validade discriminante da escala não é plena, foi estimado o modelo com oito dimensões latentes correlacionadas, o qual se mostrou superior ao modelo com oito dimensões latentes não correlacionadas. Todos os índices de adequação do ajuste desse modelo foram superiores ao anterior. No entanto, mesmo sendo este modelo superior, os índices que indicam a qualidade de ajuste não alcançaram os limites mínimos

necessários para a sua aceitação. O mediano desempenho da escala é decorrente do excesso de multicolinearidade existente entre os fatores.

O mediano desempenho da escala conduz para sua reformulação. Utilizando-se da metodologia de análise fatorial exploratória, após a eliminação dos indicadores que apresentavam mais ruídos do que informações foi diagnosticado que os discentes reconhecem consistentemente quatro dimensões correlacionadas da escala. Essas dimensões foram rotuladas como: competência, relacionamento e ética, motivação e avaliação. Todas essas dimensões estão presentes nos instrumentos de avaliação da docência dos principais instrumentos validados na literatura.

Como essa técnica é simplesmente exploratória, utilizou-se os modelos causais para confirmar as novas dimensões da escala. Os modelos foram estimados para o total de disciplinas avaliadas na UNISUL e estratificados por *campus*, que, no segundo semestre letivo de 2002, eram três.

Foram estimados e analisados oito modelos de medida de indicadores exógenos. Quatro modelos em que todos os indicadores estavam associados a um único fator e quatro modelos com quatro fatores correlacionados. Os quatro modelos de fator único evidenciaram índices de qualidade de ajuste absoluto, incremental e de parcimônia muito similares e, sobretudo, muito abaixo dos limites teóricos desejáveis. Enquanto que os modelos com quatro fatores latentes correlacionados apresentaram para os mesmos indicadores, índices superiores de ajuste. Ressaltam-se os índices de ajuste incrementais que comparam o modelo com quatro fatores latentes correlacionados com o modelo em que todos os indicadores estão associados a um único fator latente. Esse último modelo alcançou índices de ajuste incrementais dentro dos limites necessários para validar os construtos operacionalizados no instrumento, em todos os modelos com quatro fatores correlacionados estimados.

No global, os índices de qualidade de ajuste obtidos pelos modelos com quatro fatores latentes correlacionados, são menores que os obtidos pelos instrumentos validados e

citados na literatura. Contudo, o índice de modificação obtido indica que não são necessários mais ajustes.

Em síntese, os diversos índices de confiabilidade dos indicadores e dos construtos decorrentes dos estudos confirmatórios são superiores 0,7, indicando consistência nas avaliações dos discentes. 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. Certamente, o ensino é um processo muito complexo, e não pode ser resumido por essas quatro dimensões. Continuadamente, as instituições de ensino precisam adequar os seus instrumentos, incluindo e testando novos indicadores que contemplem todos os paradigmas que norteiam as diferentes estratégias de avaliação do ensino.

Quanto à validade de critério, todas as dimensões validadas da escala causam as variáveis critérios “indicaria este professor a outras turmas” e “professor conduz o ensino da disciplina com qualidade”. As altas correlações encontradas indicam que as dimensões validadas da escala se manifestam nessas variáveis “critérios”, durante o processo de avaliação de disciplinas.

Também, a maioria dos indicadores das dimensões da escala estão correlacionados positivamente e significativamente com os seguintes variáveis critérios: “considera que o seu curso tem qualidade”, “o seu curso está exigindo o suficiente para a sua formação profissional”, “considera a UNISUL uma universidade de qualidade e “você tem orgulho de estudar na UNISUL”. As correlações encontradas entre os indicadores da escala com essas variáveis critérios evidenciam a validade da escala, para medir os níveis de

satisfação dos alunos para com as diversas dimensões do ensino eficaz e sustentam a hipótese de que, se a universidade possuir um corpo docente qualificado e em atualização contínua, esta será uma condição fundamental para o êxito institucional, conseqüentemente, dos novos formados que semestralmente são inseridos no mercado de trabalho.

Quanto à capacidade da escala para diagnosticar tipologias “professores”, as diversas análises de cluster realizadas evidenciaram quatro *cluster* que possibilitam discriminar os professores em função da satisfação evidenciadas pelos discentes. A comprovação dessa hipótese acumula mais evidências de validade para o instrumento reduzido em análise.

Dessa forma, todos os indicativos de confiabilidade e de validade, diagnosticado nesse estudo de caso, indicam que as avaliações dos docentes pelos discentes sobre as características do ensino daquele professor em particular são válidas e confiáveis, portanto, os resultados provenientes do processo avaliativo são informações relevantes que podem auxiliar o professor no melhoramento contínuo da qualidade do seu ensino.

As confirmações dos atributos de confiabilidade e de validade das valorações estudantis, além de providenciar seguridade para o uso das diversas informações decorrentes do processo avaliativo, dotam o instrumento analisado de alta qualidade técnica, uma vez que confiabilidade e validade proporcionam os índices de qualidade desejáveis dos instrumentos de avaliação.

Do ponto de vista das limitações dessa 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 em 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, essa pesquisa sugere que sejam realizadas pesquisas sobre o efeito das variáveis “tamanhos das classes”, “dificuldade da disciplina”, “alunos poucos preparados para prosseguirem no curso elegido”, “idades dos alunos”, “sexo dos alunos”, “tipo de disciplinas”, “notas”, “semestre letivo” e “outras” variáveis que podem provocar vieses negativos nos níveis de satisfação dos alunos quanto ao ensino intermediado pelo professor. Recomenda, ainda, 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 validade de uso dos resultados das avaliações dos alunos, por parte dos docentes.

Além disso, a metodologia, aqui proposta 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.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMI, P. C. et al. Do teacher standards for assigning grades affect student evaluations of instruction? **Journal of Educational Psychology**, n. 72, p. 107-118.

_____. How should we use student ratings to evaluate teaching? **Research in Higher Education**, n. 30, v. 2, p. 221-227, 1989.

AKAIKE, H. Factor analysis and AIC. **Psychometrika**, n. 52, p. 317-332, 1987.

ALEAMONI, L. M. Student rating myths versus research facts from 1924 to 1998. **Journal of Personnel Evaluation in Education**, 13:2 p. 153-166, 1999.

_____. Student rating of instruction. In **J. Millman (ed), handbook of teacher evaluation**. Beverly Hills, CA: Sage, p. 110-145, 1981.

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, SP: Papirus, 2002, p. 19-190.

ANASTASI, A. Evolving concepts of test validation. **Annual Review of Psychology**, v. 37, p. 1-15, 1986.

AQUILAR, J. M.; ANDER-EGG, E. **Avaliação de serviços e programas sociais**. Petrópolis: Vozes, 1994.

ARRUDA, José Ricardo Campelo. **Políticas e indicadores da qualidade na educação superior**. Rio de Janeiro: Quality Mark/Dunya, 1997.

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 serviço n. ED 202410).

_____. **Reliability, validity and generalizability of student ratings of instruction** (IDEA paper n. 6). Manhattan: Kansas State University, Center for Faculty Evaluation and Development, 1981. (ERIC document reproduction service n. 213196).

BARTON, R. M. et al. Factorial validity and reliability of a survey to assess the teaching effectiveness of graduates of teacher education programs. **Educational and Psychological Measurement**, v. 54, n. 1, p. 218-226, spring 1994.

BATISTA – FOGUET, J. M. et al. Análisis Factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. **Med Clin**, Barcelona, v. 122, n.1, p.21-27, 2004.

_____. Factorial validity and reliability of a survey to assess the teaching effectiveness of graduates of teacher education programs. **Educational and Psychological Measurement**, v. 54, n. 1, p. 218-226, 1994.

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

_____. Significance tests and goodness of fit the analysis of covariance structures. **Psychological Bulletin**, n. 88, p. 588-606, 1980.

BLANCO, L. S. & PAZ, J. J. SANZ. Evaluacion de la docencia mediante cuestionario de alumnos: Universidade de Cantabria (curso 86-87). **Studia Paedagogica**, v. 20, p. 41-71, 1988.

BURDSAL, C. A. & BARDO, J. W. Measuring student's perceptions of teaching: dimensions of evaluations. **Educational and Psychological Measurement**, v. 46, p. 63-79, 1986.

CAMPBELL, D.T. & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. **Psychological Bulletin**, n 56, p. 81-105, 1959.

CANDE, J. et al. **Avaliação da actividade docente do profesorado universitario pelos/as alumnos/as da Universidade de Santiago**. Informe global. Universidade de Santiago, 1992, ICE.

CAPELLERAS, Joan-Lluíz, y Veciana, J. M^a. **Calidade de servico en la enseñanza universitária**: desarrollo y validación de uma escala de medida. XI Congresso Nacional de ACEDE, 2000. Disponível em: <www.fnac.es>. Acesso em 18/09/2003

CARVALHO, Eduardo Búrigo de. et al. **A avaliação institucional**: um processo permanente. Tubarão: Editora Universitária UNISUL, 1999.

CASANUEVA, C., PERIÁÑEZ, R. & J. E. RUFINO. **Calidad percibida por el alumno en el servicio docente universitario**: desarrollo de una escala de medida. XI Congresso Nacional AEDEM Lleida, v. 2, p. 27-34, 1997.

CASHIN, W. E. **Student ratings of teaching**: a summary of the research. IDEA paper n. 20. Manhattan, K. S: Kansas State University, Center for Faculty Evaluation and Development, 1988. (ERIC document reproduction service n. ED 302567).

_____. **Student ratings of teaching**: a summary of the research. IDEA paper n. 32. Manhattan, K. S: Kansas State University, Center for Faculty Evaluation and Development, 1995.

CASTRO, Cláudio de Moura. **E quem avalia os professores?** Dois pontos: teoria e prática em educação, Belo Horizonte, v.2, n. 13, p 12-13, 1992.

CENTRA, J. A. Student ratings of instruction and their relationship to student learning. **American Educational Research Journal**, v. 14, n. 1, p. 17-24, 1977.

CHURCHILL, G. A. A paradigm for developing better measures of marketing constructs. **Journal of Marketing Research**, n. 16, p. 64-74, 1979.

COELHO, J. **A avaliação institucional como instrumento de identificação de perdas acadêmicas no processo ensino-aprendizagem**. (Tese de doutorado) - Universidade Federal de Santa Catarina, 2003.

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

CONFERÊNCIA MUNDIAL SOBRE O ENSINO SUPERIOR, 1998, Paris. **Tendências da Educação Superior para o século XXI**. Brasília: UNESCO, 1999 p. 726.

CONSELHO DE REITORES DAS UNIVERSIDADES BRASILEIRAS. **Proposta de Modelo de Avaliação Institucional para as Universidades Brasileiras**. Brasília, 2000.

COSTIN, F. et ali. Student ratings of College Teaching: reliability, validity, and usefulness. **Review of Educational Research**, 1971. p. 511-535.

COVERT, R. W & MASON, E. J. Factorial validity of a student evaluation of teaching instrument. **Educational and psychological Measurement**, n. 34, p. 903-905, 1974.

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

D'APOLLONIA, S. & ABRAMI, P. C. Navigating student ratings of instruction. **American Psychologist**, v. 52, n. 11, p. 1198-1208, 1997.

DARLING-HAMMOND, L. et al. Teacher evaluation in the organizational context: a review of the literature. **Review of educational research**, 53, 3, p 285-328, 1983.

DUARTE, Cláudio Manoel da Cunha. Et al. Avaliação institucional: um processo em consolidação na Universidade Católica de Pelotas. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação Institucional integrada: os dez anos do Paiung**. Ijuí: Ed. Unijuí, 2003. p 195-219.

FELDMAN, K. A. Class size and college student's evaluations of teachers and courses: a closer look. **Research in Higher Education**, v. 21, n. 1, p 45-116, 1984.

_____. Class size and college students evaluations of teachers and courses: a closes look. **Research in Higher Education**, v. 21, n. 1, p. 45-116, 1984.

_____. The association between student ratings of specific instructional dimensions and student achievement: refining and extending the synthesis of data from multisection validity studies. **Research in Higher Education**, n. 30, p. 583-645, 1989.

_____. The significance os circumstances for college student's ratings of their teachers and courses. **Research in Higher Education**, 10, p. 149-172, 1979.

FELDT, L. S.; WOODRUFF, D. J. & SALIH, F. A. Statistical inference for coefficient alpha. **Applied Psychological Measurement**. n. 11, v.1, p. 93-103, 1987.

FRENCH-LAZOVIK, G. Predictability of students' evaluations of college teachers from component ratings. **Journal of Educational Psychology**, v. 66, n. 3, p. 373-385, 1974.

FREY, P. W., LEONARD, D. W & BEATTY, W. W. Student ratings of instruction: validation researsh. **American Educational Research Journal**, v. 12, n. 4, p 435-447, 1975.

- GARCIA RAMOS, J. M. e GONGOSTO LUNA, E. “**Evaluación e qualidade del professorado**”. In GONZÁLEZ RAMÍREZ, T. (Org.). *Evaluación y Gestión de la Calidad Educativa. Un enfoque Metodológico*. Málaga: Aljibe, 2000. p. 127-157.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 1994.
- GREENWALD, A. G. & GILLMORE, G. M. **Grading leniency is a removable contaminant of student ratings**. *American Psychologist*, v. 52, n. 11, p. 1209-1217, 1997.
- GREENWALD, A. G. Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, v. 52, n. 11, p. 1182-1186, 1997.
- 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*, v. 28, n. 3, p. 229-238, 2003.
- HAIR, J. F. et al. **Análises multivariantes**. 5. ed. Madrid: Prentice Hall Iberia, 1999.
- HILDEBRAND, M., WILSON, R. C & DIENST, E.R. **Evaluating university teaching**. Berkeley: Center for Research and Development in Higher Education, University of California, Berkeley, 1971.
- HOBSON, S. M. & TALBOT, D. M. **Understanding student evaluations**. *College Teaching*, n. 49, v. 1, p. 26-32.
- HOWARD, G. S. & MAXWELL, S. E. Do grades contaminate student evaluations of instruction? *Research in Higher Education*, n. 16, p. 175-188.
- HOWARD, G. S; CONWAY, C. G. Construct validity of measures of college teaching effectiveness. *Journal of Educational Psychology*, v. 77, n. 2, p 187-196, 1985.
- JOHNSON, R. A. & WICHERN, D. W. **Applied multivariate statistical analysis**. New Jersey: Prentice Hall, 1998.
- JÖRESKOG, K. G. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, n. 34, p. 183-202, 1969.
- _____. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, n. 34, p. 183-202, 1969.
- _____. Structural analysis of covariance and correlation matrices. *Psychometrika*, v. 43, n. 4, p. 443-477, 1978.
- KAISER, H. F. “**An index of factorial simplicity**”. *Psychometrika*, n. 30, p. 1-14, 1974.
- KIM, J. O. & MUELLER, C. W. **Introduction to factor analysis**. Beverly Hills: Sage Publications, 1978.
- KLINE, P. **An easy guide to factor analysis**. London: Rontledge, 1994.
- LEHMKUHL, José Mauro (Org.). **Avaliação institucional**. Joaçaba: UNOESC, 2002.

LI, R. & KAYE, M. A case study for comparing two service quality measurement approaches in the context of teaching in higher education. **Quality in Higher Education**, v. 4, n. 2, p. 103-113, 1998.

MACEDO, Sandré Granzotto. **Desempenho docente pela avaliação discente**: uma proposta metodológica para subsidiar a gestão universitária. (Tese de doutorado) - Universidade Federal de Santa Catarina, Florianópolis, 2001.

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

MARCONI, M. A. LAKATOS, E. M. **Técnicas de pesquisa**. São Paulo: Atlas 2002.

MARSH, H. W, & DUNKIN, M. **Student's evaluations of university teaching**: a multidimensional perspective. In J.C. Smart (ed) *Higher Education: Handbook of Theory and research*. Vol. 8. p. 143-233. New York: Agathon, 1992.

MARSH, H. W. & HOCEVAR, D. Confirmatory factor analysis of multitrait-multimethod matrices. **Journal of Educational Measurement**, 1983. p. 231-248.

_____. 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.

_____. The multidimensionality of students' evaluations of teaching effectiveness: the generality of factor structures across academic discipline, instructor level, and course level. **Teaching & Teacher Education**, v. 7, n. 1, p. 9-18, 1991.

MARSH, H. W. & ROCHE, L. A. Making student's evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. **American Psychologist**, 52, p. 1187-1197, 1997.

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

MARSH, H. W. Multidimensional student's evaluations of teaching effectiveness: a test of alternative higher-order structures. **Journal Educational Psychology**, v. 83, n. 2, p. 285-296, 1991.

_____. Students evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. **Journal of Educational Psychology**, v. 76, n. 5, p 707-754, 1984.

_____. Validity of student's evaluations of college teaching: a multitrait-multimethod analysis. **Journal of Educational Psychology**, v. 74, p. 264-279, 1982.

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

MARSH, H.W. & HOCEVAR, D. The multidimensionality of student's evaluations of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. **Teaching & Teacher Education**, v. 7, n. 1, p 9-18, 1991.

- MARTÍNEZ, T. L. Análisis factorial. In: MARTÍNEZ, Teodoro Luque (Org.). **Técnicas de análisis de datos en investigación de mercados**. Madrid: Pirámide, 2000.
- 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, 1992, n. 11, Madrid: Editorial Complutense, p. 73-80.
- MCDONALD, R. P. & MARSH, H. W. Choosing a multivariate model: noncentrality and goodness of fit. **Psychological Bulletin**, n. 107, p. 247-55, 1990.
- MCKEACHIE, W. J. Student ratings of faculty: a reprise. **Academe**, n. 65, p. 384-397, 1979.
- _____. Student ratings: the validity of use. **American Psychologist**, v. 52, n. 11, p. 1218-1225, 1997.
- MOREIRA, Daniel A. **Avaliação do professor universitário pelo aluno: possibilidades e limitações**. (Tese de doutorado) -Universidade de São Paulo, 1986.
- MULAİK, S. A. et al. Evaluation of goodness-of-fit indexes for structural equation models. **Psychological Bulletin**, v. 105, n. 3, p. 340-445, 1989.
- _____. Evaluation of goodness-of-bit indexes for structural equation models. **Psychological Bulletin**, v. 105, n. 3, p 430-445, 1989.
- NARDIA, K. V. Measures of Multivariate Shewness and kurtosis with applications, **Biometrika** v.57, pp. 519-530, 1970.
- NASCIMENTO, A. F. M. **Avaliação de programas, projetos e atividades universitárias: referenciando a prática**. Roberto Lassance, Brasília: Universa, 2000. p. 9-61.
- NAUJORKS, Luiz Carlos et al. A avaliação institucional na Universidade de Passo Fundo no contexto dos dez anos do PAIUNG. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG**. Ijuí: Ed. Unijuí, 2003. p. 93-113.
- NEWALL, D. & DALE, B. G. Measuring quality improvement: a management critique. **Total quality management**, 2, p. 255-267, 1991.
- NUNES, Ana Karin. A Avaliação Institucional na UNISC: um processo transformador moldado pelos pressupostos fundamentais do PAIUNG. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG**. Ijuí: Unijuí, 2003. p. 69-91.
- NUNNALLY, J. C. **Psychometric Theory**. New York: McGraw-Hill, 1978.
- OVERALL, J. V. & MARSH, H. W. Student's evaluations of instruction: a longitudinal study os their stability. **Journal of Educational Psychology**, n. 72, p. 321-325, 1980.
- _____. Student's evaluations of instruction: a longitudinal study of their stability. **Journal Of Educational Psychology**, 72, p. 321-325, 1980.

OWLIA, M. S. & ASPINWALL, E. M. A framework for measuring quality in engineering education. **Total Quality Management**, v. 9, n. 6, p. 501-518, 1998.

_____. **Quality in higher education: a survey total quality management**, 7 (4), p. 161-171, 1996.

PASQUALI, Luiz. Questionário de avaliação de docência. **Educação e Seleção**, São Paulo, n. 9, p. 71-98, 1984.

PEREIRA, J. C. R. **Análise de dados qualitativos**. Estratégias metodológicas para as ciências da saúde, humanas e sociais. São Paulo: Editora da Universidade de São Paulo, 1999. p. 1-154.

PÉREZ, M. S.; SÁNCHEZ, F. J. S. Validez y fiabilidad de escala. In: SÁNCHEZ, Francisco José Saraiba (Org.). **Metodología para la investigación en marketing y dirección de empresas**. Madrid: Pirámide, 1999. p. 362-388.

PRESTES, Gelça Regina Lusa. Et al. Avaliação da UCS. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG**. Ijuí: Unijuí, 2003. p. 139-151.

RAJU, N. S. A genelization of coefficient alpha. **Psychometrika**, v. 42, p. 549-565, 1977.

RAMOS, Inajara Vargas. Avaliação Institucional: a inserção da FEEVALE. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG**. Ijuí: Unijuí, 2003. p. 221-230.

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

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

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

RAMSDEN, P. A performance indicator of teaching quality in higher education. The experience questionnaire. **Studies in Higher Education**, n. 16, p. 129-150, 1991.

REMMERS, H. H. et al. Are student rating of their instructors related to their grades? **Purdue University Studies in Higher Education**, n. 44, p. 17-26, 1949.

REQUENA, C. S. **Psicometria**. Teoria y practica en la construccion de tests. Ediciones Norma, Madrid, 1990.

RÖSLER, Mara Regina et al. Uma reflexão necessária sobre a trajetória da avaliação na URI. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG**. Ijuí: Unijuí, 2003. p. 195-219.

SANCHEZ, A. V. La evaluación del professor: perspectivas y resultados. **Revista de Educación**, v. 27, p. 55-93, 1985.

SÁNCHEZ, J. F. **La evaluación de la enseñanza universitaria: la experiencia de La Complutense.** *Studia Paedagogica*, 1988, v. 20, p. 135-146.

SHAH, A. et al. **The influence of bias factors on student rating of teaching.** City University of Hong Kong, 2002. Disponível em <http://succeednow.org/icee/Presentations/Pres_340_468.pdf> Acesso em 20/03/2004

SHERMAN, B. R. & BLACKBURN, R. T. Personal characteristics and teaching effectiveness of college faculty. **Journal of Educational Psychology**, v. 57, n. 1, p. 124-131, 1975.

_____. Personal characteristics and teaching effectiveness of college faculty. **Journal of Educational Psychology**, v. 1, p. 124-131, 1975.

SILUK, Ana Cláudia Pavão et al. Avaliação Institucional na UNICRUZ: um processo de reconstrução. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG.** Ijuí: Unijuí, 2003. p. 139-151.

SILVA, Ilton Benoni da. A avaliação como prática intrínseca ao fazer universitário: a experiência da Unijuí. In: SILVA, Ilton Benoni da; ROSA, Magna Stella Cargnelutti Dalla. (Org.). **Avaliação institucional integrada: os dez anos do PAIUNG.** Ijuí: Unijuí, 2003. p. 115-137.

SILVEIRA, F. L. da. et al. Avaliação do desempenho do professor pelo aluno: novas evidências de validade de um instrumento. **Ciência e Cultura**, n. 37, v. 8, p. 1237-1240, 1985.

SILVEIRA, F. L. da e MOREIRA, M. A. Avaliação do desempenho do professor pelo aluno: evidências de validade de um instrumento. **Ciência e Cultura**, n. 36, p. 466-472, 1984.

SNYDER, C. R & CLAIR, M. Effects of expected and obtained grades on teacher evaluation and attribution of performance. **Journal of Educational Psychologist**, n. 68, p. 75-82, 1976.

SOBRINHO, José Dias. Educação e Avaliação: Técnica e Ética. In: _____; RISTOFF, Dilvo I. (Org.). **Avaliação democrática: para uma universidade cidadã.** Florianópolis: Insular, 2002. p. 37-38.

STAPLETON, C. D. Basic concepts in exploratory factor analysis (EFA) as a tool to evaluate score validity: A right-brained approach. **Texas A & M University**, 1997. Disponível em: <<http://erical.net/ft/tamu/efa.htm>>. Acesso em: 31 maio 2002.

SUCH, J. G. et al. **Análises de tipologías de calidad docente a partir de um cuestionario de evaluación del profesorado universitario.** *Bordon*, n. 51, v.1, p. 95-113, 1999.

SUCH, José González. **Algumas notas sobre a docência Universitat de València.** Disponível em: <<http://www.ub.es/forum/conferencias/such.htm>>. Acesso em: 27 mar. 2003.

TEJEDOR, F. J. & GARCIA VALCÁRCEL, A. “Evaluación institucional en la universidad”. **Revista Galega de Psicopedagogia**. v. 6, p. 101-146, 1996.

TEJEDOR, F. J. & MONTEIRO, M^a. L. Indicadores de la calidad docente para la evaluación del profesor universitario. **Revista Española de Pedagogía**, año XLVIII, n. 186, p. 259-279, mayo-agosto, 1990.

_____. Indicadores de la calidad docente para la evaluación del profesor universitario. **Revista Española de Pedagogía**, n. 186, p. 259-279, 1990.

TEJEDOR, F. J. et al. Evaluación del profesorado universitario por los alumnos en la universidad de Santiago. **Studia Paedagogica**, v. 20, p. 73-134, 1988.

TINSLEY, H. E. A. & TINSLEY, D. J. Uses of factor analysis in counseling psychology research. **Journal of Counseling Psychology**, n. 34, p. 414-424.

TOURÓN, J. **La validación de constructo**: sua aplicación al CEED (Cuestionario para la evaluación de la eficacia docente. Bordón, 1989. p. 735-756.

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

UMBLE, M. M. & WHITTEN, B. J. The significant dimensions of teaching behavior and their relative importance for instructor evaluations. **Educational and Psychological Measurement**, v. 37, p. 1023-1030, 1977.

_____. The significant dimensions of teaching behavior and their relative importance for instructor evaluations. **Educational and Psychological Measurement**, n. 37, p. 1023-1030, 1977.

UNIVERSIDADE DO OESTE DE SANTA CATARINA. **Cadernos de avaliação**. São Miguel.

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

VANLEEUEWEN et al. Assessing the reliability of student evaluations of teaching (SETS) with generalizability theory. **Journal of Agricultural Education**, v. 40, n. 4, p 1-9, 1999.

WHITELY, S. E & DOYLE, K. E. **"Dimensions of effective teaching: factors or artifacts"**. **Educational and Psychological Measurement**, v. 38, p. 107-117, 1978.

WORTHINGTON, A. G. & WONG, P. T. P. Effects of earned and assigned grades on student evaluations of an instructor. **Journal of Educational Psychologist**, n. 71, p. 764-775.

ANEXOS

Anexo I: Questionário Discente

nexo II: Matriz de correlação geral – Unisul

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q34	1,00	0,70	0,77	0,68	0,78	0,60	0,44	0,45	0,73	0,60	0,63	0,55	0,58	0,64	0,81	0,85	0,54	0,57	0,51	0,64	0,66	0,60	0,52	0,45	0,73	0,54
Q36	0,70	1,00	0,63	0,55	0,57	0,49	0,36	0,33	0,55	0,46	0,55	0,49	0,56	0,60	0,68	0,64	0,46	0,50	0,43	0,57	0,54	0,56	0,49	0,43	0,62	0,47
Q37	0,77	0,63	1,00	0,70	0,80	0,55	0,41	0,43	0,68	0,62	0,65	0,56	0,58	0,64	0,78	0,82	0,56	0,58	0,50	0,66	0,63	0,59	0,51	0,48	0,75	0,54
Q38	0,68	0,55	0,70	1,00	0,78	0,50	0,41	0,43	0,84	0,65	0,73	0,70	0,62	0,71	0,66	0,78	0,71	0,67	0,64	0,61	0,64	0,59	0,53	0,61	0,76	0,67
Q39	0,78	0,57	0,80	0,78	1,00	0,65	0,48	0,52	0,79	0,71	0,69	0,61	0,59	0,66	0,78	0,86	0,66	0,65	0,61	0,68	0,70	0,62	0,55	0,53	0,78	0,61
Q40	0,60	0,49	0,55	0,50	0,65	1,00	0,65	0,74	0,59	0,65	0,50	0,40	0,47	0,49	0,63	0,62	0,42	0,54	0,47	0,59	0,61	0,57	0,51	0,36	0,56	0,49
Q41	0,44	0,36	0,41	0,41	0,48	0,65	1,00	0,72	0,51	0,59	0,38	0,37	0,43	0,44	0,47	0,46	0,36	0,52	0,47	0,54	0,50	0,52	0,51	0,33	0,46	0,49
Q42	0,45	0,33	0,43	0,43	0,52	0,74	0,72	1,00	0,53	0,69	0,42	0,34	0,39	0,39	0,47	0,50	0,35	0,54	0,48	0,51	0,53	0,50	0,48	0,32	0,47	0,48
Q43	0,73	0,55	0,68	0,84	0,79	0,59	0,51	0,53	1,00	0,73	0,73	0,71	0,66	0,73	0,71	0,82	0,70	0,70	0,70	0,65	0,70	0,64	0,58	0,60	0,77	0,71
Q44	0,60	0,46	0,62	0,65	0,71	0,65	0,59	0,69	0,73	1,00	0,67	0,55	0,57	0,60	0,63	0,71	0,61	0,73	0,67	0,63	0,63	0,62	0,59	0,55	0,69	0,68
Q45	0,63	0,55	0,65	0,73	0,69	0,50	0,38	0,42	0,73	0,67	1,00	0,67	0,67	0,69	0,65	0,74	0,80	0,73	0,69	0,61	0,64	0,63	0,54	0,75	0,79	0,72
Q46	0,55	0,49	0,56	0,70	0,61	0,40	0,37	0,34	0,71	0,55	0,67	1,00	0,76	0,82	0,58	0,65	0,67	0,62	0,61	0,53	0,54	0,54	0,50	0,61	0,70	0,64
Q47	0,58	0,56	0,58	0,62	0,59	0,47	0,43	0,39	0,66	0,57	0,67	0,76	1,00	0,78	0,63	0,66	0,66	0,63	0,62	0,58	0,58	0,61	0,57	0,62	0,70	0,64
Q48	0,64	0,60	0,64	0,71	0,66	0,49	0,44	0,39	0,73	0,60	0,69	0,82	0,78	1,00	0,71	0,74	0,66	0,65	0,61	0,63	0,62	0,62	0,58	0,62	0,74	0,65
Q49	0,81	0,68	0,78	0,66	0,78	0,63	0,47	0,47	0,71	0,63	0,65	0,58	0,63	0,71	1,00	0,86	0,55	0,62	0,54	0,68	0,70	0,65	0,57	0,47	0,75	0,57
Q50	0,85	0,64	0,82	0,78	0,86	0,62	0,46	0,50	0,82	0,71	0,74	0,65	0,66	0,74	0,86	1,00	0,68	0,69	0,63	0,71	0,73	0,69	0,59	0,58	0,84	0,66
Q51	0,54	0,46	0,56	0,71	0,66	0,42	0,36	0,35	0,70	0,61	0,80	0,67	0,66	0,66	0,55	0,68	1,00	0,72	0,78	0,58	0,60	0,61	0,54	0,84	0,77	0,76
Q52	0,57	0,50	0,58	0,67	0,65	0,54	0,52	0,54	0,70	0,73	0,73	0,62	0,63	0,65	0,62	0,69	0,72	1,00	0,78	0,66	0,66	0,70	0,62	0,68	0,75	0,79
Q53	0,51	0,43	0,50	0,64	0,61	0,47	0,47	0,48	0,70	0,67	0,69	0,61	0,62	0,61	0,54	0,63	0,78	0,78	1,00	0,66	0,64	0,68	0,65	0,71	0,72	0,84
Q54	0,64	0,57	0,66	0,61	0,68	0,59	0,54	0,51	0,65	0,63	0,61	0,53	0,58	0,63	0,68	0,71	0,58	0,66	0,66	1,00	0,77	0,75	0,76	0,52	0,72	0,65
Q55	0,66	0,54	0,63	0,64	0,70	0,61	0,50	0,53	0,70	0,63	0,64	0,54	0,58	0,62	0,70	0,73	0,60	0,66	0,64	0,77	1,00	0,73	0,67	0,51	0,71	0,65
Q56	0,60	0,56	0,59	0,59	0,62	0,57	0,52	0,50	0,64	0,62	0,63	0,54	0,61	0,62	0,65	0,69	0,61	0,70	0,68	0,75	0,73	1,00	0,82	0,61	0,72	0,71
Q57	0,52	0,49	0,51	0,53	0,55	0,51	0,51	0,48	0,58	0,59	0,54	0,50	0,57	0,58	0,57	0,59	0,54	0,62	0,65	0,76	0,67	0,82	1,00	0,54	0,64	0,66
Q58	0,45	0,43	0,48	0,61	0,53	0,36	0,33	0,32	0,60	0,55	0,75	0,61	0,62	0,62	0,47	0,58	0,84	0,68	0,71	0,52	0,51	0,61	0,54	1,00	0,73	0,75
Q59	0,73	0,62	0,75	0,76	0,78	0,56	0,46	0,47	0,77	0,69	0,79	0,70	0,70	0,74	0,75	0,84	0,77	0,75	0,72	0,72	0,71	0,72	0,64	0,73	1,00	0,77
Q60	0,54	0,47	0,54	0,67	0,61	0,49	0,49	0,48	0,71	0,68	0,72	0,64	0,64	0,65	0,57	0,66	0,76	0,79	0,84	0,65	0,65	0,71	0,66	0,75	0,77	1,00

Anexo III: Matrix de correlação campus Tubarão

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q34	1,00	0,67	0,77	0,68	0,79	0,58	0,36	0,41	0,70	0,59	0,63	0,50	0,53	0,62	0,82	0,85	0,51	0,55	0,46	0,62	0,66	0,54	0,47	0,39	0,72	0,47
Q36	0,67	1,00	0,62	0,53	0,55	0,51	0,33	0,31	0,54	0,47	0,54	0,44	0,55	0,59	0,69	0,64	0,44	0,49	0,42	0,58	0,55	0,56	0,48	0,39	0,62	0,44
Q37	0,77	0,62	1,00	0,70	0,80	0,52	0,35	0,38	0,67	0,60	0,65	0,51	0,56	0,61	0,77	0,82	0,54	0,57	0,47	0,62	0,63	0,57	0,47	0,43	0,75	0,48
Q38	0,68	0,53	0,70	1,00	0,76	0,45	0,36	0,39	0,83	0,65	0,73	0,67	0,58	0,68	0,66	0,79	0,70	0,65	0,63	0,57	0,63	0,56	0,47	0,57	0,75	0,64
Q39	0,79	0,55	0,80	0,76	1,00	0,63	0,41	0,48	0,77	0,68	0,67	0,55	0,54	0,63	0,78	0,86	0,63	0,63	0,56	0,65	0,71	0,57	0,48	0,46	0,77	0,54
Q40	0,58	0,51	0,52	0,45	0,63	1,00	0,61	0,72	0,56	0,60	0,45	0,33	0,41	0,46	0,62	0,61	0,36	0,50	0,45	0,59	0,61	0,53	0,48	0,27	0,52	0,42
Q41	0,36	0,33	0,35	0,36	0,41	0,61	1,00	0,71	0,48	0,54	0,32	0,31	0,42	0,42	0,41	0,41	0,29	0,50	0,47	0,54	0,44	0,52	0,54	0,27	0,40	0,46
Q42	0,41	0,31	0,38	0,39	0,48	0,72	0,71	1,00	0,52	0,68	0,37	0,28	0,33	0,36	0,43	0,48	0,30	0,52	0,46	0,50	0,50	0,47	0,45	0,24	0,42	0,45
Q43	0,70	0,54	0,67	0,83	0,77	0,56	0,48	0,52	1,00	0,73	0,73	0,70	0,63	0,72	0,70	0,81	0,69	0,70	0,68	0,63	0,69	0,60	0,53	0,56	0,77	0,69
Q44	0,59	0,47	0,60	0,65	0,68	0,60	0,54	0,68	0,73	1,00	0,68	0,52	0,54	0,59	0,62	0,71	0,60	0,74	0,66	0,61	0,62	0,61	0,56	0,50	0,68	0,67
Q45	0,63	0,54	0,65	0,73	0,67	0,45	0,32	0,37	0,73	0,68	1,00	0,64	0,62	0,67	0,63	0,73	0,78	0,70	0,64	0,56	0,60	0,58	0,47	0,70	0,77	0,67
Q46	0,50	0,44	0,51	0,67	0,55	0,33	0,31	0,28	0,70	0,52	0,64	1,00	0,74	0,81	0,52	0,61	0,66	0,60	0,60	0,47	0,47	0,50	0,45	0,58	0,65	0,62
Q47	0,53	0,55	0,56	0,58	0,54	0,41	0,42	0,33	0,63	0,54	0,62	0,74	1,00	0,79	0,56	0,61	0,62	0,58	0,56	0,53	0,49	0,56	0,50	0,57	0,64	0,58
Q48	0,62	0,59	0,61	0,68	0,63	0,46	0,42	0,36	0,72	0,59	0,67	0,81	0,79	1,00	0,66	0,71	0,65	0,64	0,60	0,60	0,57	0,60	0,55	0,59	0,71	0,63
Q49	0,82	0,69	0,77	0,66	0,78	0,62	0,41	0,43	0,70	0,62	0,63	0,52	0,56	0,66	1,00	0,87	0,51	0,60	0,51	0,67	0,70	0,62	0,53	0,39	0,75	0,52
Q50	0,85	0,64	0,82	0,79	0,86	0,61	0,41	0,48	0,81	0,71	0,73	0,61	0,61	0,71	0,87	1,00	0,66	0,68	0,60	0,69	0,73	0,66	0,55	0,52	0,83	0,62
Q51	0,51	0,44	0,54	0,70	0,63	0,36	0,29	0,30	0,69	0,60	0,78	0,66	0,62	0,65	0,51	0,66	1,00	0,70	0,74	0,54	0,54	0,56	0,49	0,81	0,75	0,72
Q52	0,55	0,49	0,57	0,65	0,63	0,50	0,50	0,52	0,70	0,74	0,70	0,60	0,58	0,64	0,60	0,68	0,70	1,00	0,77	0,62	0,63	0,68	0,57	0,65	0,72	0,78
Q53	0,46	0,42	0,47	0,63	0,56	0,45	0,47	0,46	0,68	0,66	0,64	0,60	0,56	0,60	0,51	0,60	0,74	0,77	1,00	0,63	0,57	0,65	0,63	0,67	0,67	0,83
Q54	0,62	0,58	0,62	0,57	0,65	0,59	0,54	0,50	0,63	0,61	0,56	0,47	0,53	0,60	0,67	0,69	0,54	0,62	0,63	1,00	0,76	0,76	0,77	0,45	0,68	0,59
Q55	0,66	0,55	0,63	0,63	0,71	0,61	0,44	0,50	0,69	0,62	0,60	0,47	0,49	0,57	0,70	0,73	0,54	0,63	0,57	0,76	1,00	0,70	0,61	0,42	0,69	0,57
Q56	0,54	0,56	0,57	0,56	0,57	0,53	0,52	0,47	0,60	0,61	0,58	0,50	0,56	0,60	0,62	0,66	0,56	0,68	0,65	0,76	0,70	1,00	0,80	0,56	0,71	0,69
Q57	0,47	0,48	0,47	0,47	0,48	0,48	0,54	0,45	0,53	0,56	0,47	0,45	0,50	0,55	0,53	0,55	0,49	0,57	0,63	0,77	0,61	0,80	1,00	0,47	0,59	0,62
Q58	0,39	0,39	0,43	0,57	0,46	0,27	0,27	0,24	0,56	0,50	0,70	0,58	0,57	0,59	0,39	0,52	0,81	0,65	0,67	0,45	0,42	0,56	0,47	1,00	0,68	0,71
Q59	0,72	0,62	0,75	0,75	0,77	0,52	0,40	0,42	0,77	0,68	0,77	0,65	0,64	0,71	0,75	0,83	0,75	0,72	0,67	0,68	0,69	0,71	0,59	0,68	1,00	0,73
Q60	0,47	0,44	0,48	0,64	0,54	0,42	0,46	0,45	0,69	0,67	0,67	0,62	0,58	0,63	0,52	0,62	0,72	0,78	0,83	0,59	0,57	0,69	0,62	0,71	0,73	1,00

Anexo IV: Matrix de correlação Campus Araranguá

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q34	1,00	0,71	0,79	0,67	0,77	0,56	0,42	0,45	0,70	0,58	0,60	0,55	0,55	0,65	0,80	0,82	0,54	0,56	0,53	0,67	0,64	0,62	0,55	0,47	0,73	0,55
Q36	0,71	1,00	0,61	0,51	0,55	0,47	0,39	0,38	0,54	0,45	0,53	0,47	0,55	0,58	0,66	0,64	0,47	0,53	0,44	0,57	0,55	0,63	0,55	0,46	0,60	0,51
Q37	0,79	0,61	1,00	0,71	0,84	0,57	0,36	0,44	0,71	0,64	0,67	0,58	0,61	0,67	0,81	0,85	0,59	0,62	0,54	0,71	0,64	0,63	0,57	0,51	0,79	0,58
Q38	0,67	0,51	0,71	1,00	0,79	0,54	0,43	0,45	0,87	0,67	0,77	0,73	0,66	0,75	0,66	0,75	0,75	0,73	0,71	0,70	0,66	0,66	0,62	0,68	0,80	0,73
Q39	0,77	0,55	0,84	0,79	1,00	0,64	0,48	0,51	0,79	0,71	0,69	0,64	0,63	0,70	0,77	0,86	0,70	0,69	0,63	0,73	0,68	0,66	0,62	0,62	0,81	0,66
Q40	0,56	0,47	0,57	0,54	0,64	1,00	0,67	0,77	0,64	0,71	0,53	0,48	0,52	0,55	0,63	0,60	0,48	0,61	0,56	0,67	0,66	0,66	0,59	0,44	0,63	0,59
Q41	0,42	0,39	0,36	0,43	0,48	0,67	1,00	0,66	0,56	0,58	0,43	0,41	0,43	0,43	0,46	0,46	0,45	0,56	0,56	0,55	0,58	0,53	0,49	0,41	0,53	0,60
Q42	0,45	0,38	0,44	0,45	0,51	0,77	0,66	1,00	0,56	0,67	0,44	0,44	0,44	0,44	0,47	0,47	0,40	0,57	0,52	0,55	0,56	0,54	0,51	0,40	0,53	0,53
Q43	0,70	0,54	0,71	0,87	0,79	0,64	0,56	0,56	1,00	0,75	0,77	0,77	0,71	0,79	0,71	0,79	0,76	0,76	0,75	0,73	0,72	0,73	0,67	0,72	0,83	0,79
Q44	0,58	0,45	0,64	0,67	0,71	0,71	0,58	0,67	0,75	1,00	0,70	0,61	0,65	0,68	0,68	0,71	0,63	0,74	0,67	0,70	0,69	0,68	0,68	0,62	0,76	0,71
Q45	0,60	0,53	0,67	0,77	0,69	0,53	0,43	0,44	0,77	0,70	1,00	0,71	0,71	0,75	0,68	0,72	0,81	0,76	0,69	0,65	0,65	0,69	0,63	0,79	0,80	0,75
Q46	0,55	0,47	0,58	0,73	0,64	0,48	0,41	0,44	0,77	0,61	0,71	1,00	0,81	0,86	0,59	0,65	0,67	0,69	0,63	0,62	0,64	0,63	0,62	0,68	0,74	0,69
Q47	0,55	0,55	0,61	0,66	0,63	0,52	0,43	0,44	0,71	0,65	0,71	0,81	1,00	0,82	0,65	0,65	0,67	0,72	0,63	0,68	0,68	0,70	0,70	0,71	0,75	0,74
Q48	0,65	0,58	0,67	0,75	0,70	0,55	0,43	0,44	0,79	0,68	0,75	0,86	0,82	1,00	0,72	0,77	0,70	0,71	0,61	0,71	0,73	0,73	0,69	0,70	0,81	0,72
Q49	0,80	0,66	0,81	0,66	0,77	0,63	0,46	0,47	0,71	0,68	0,68	0,59	0,65	0,72	1,00	0,85	0,57	0,64	0,54	0,74	0,68	0,70	0,63	0,54	0,77	0,59
Q50	0,82	0,64	0,85	0,75	0,86	0,60	0,46	0,47	0,79	0,71	0,72	0,65	0,65	0,77	0,85	1,00	0,67	0,68	0,62	0,78	0,72	0,74	0,67	0,60	0,85	0,67
Q51	0,54	0,47	0,59	0,75	0,70	0,48	0,45	0,40	0,76	0,63	0,81	0,67	0,67	0,70	0,57	0,67	1,00	0,74	0,78	0,61	0,59	0,64	0,60	0,88	0,77	0,81
Q52	0,56	0,53	0,62	0,73	0,69	0,61	0,56	0,57	0,76	0,74	0,76	0,69	0,72	0,71	0,64	0,68	0,74	1,00	0,78	0,71	0,72	0,76	0,71	0,73	0,80	0,83
Q53	0,53	0,44	0,54	0,71	0,63	0,56	0,56	0,52	0,75	0,67	0,69	0,63	0,63	0,61	0,54	0,62	0,78	0,78	1,00	0,67	0,68	0,69	0,70	0,72	0,70	0,86
Q54	0,67	0,57	0,71	0,70	0,73	0,67	0,55	0,55	0,73	0,70	0,65	0,62	0,68	0,71	0,74	0,78	0,61	0,71	0,67	1,00	0,84	0,79	0,78	0,59	0,76	0,73
Q55	0,64	0,55	0,64	0,66	0,68	0,66	0,58	0,56	0,72	0,69	0,65	0,64	0,68	0,73	0,68	0,72	0,59	0,72	0,68	0,84	1,00	0,80	0,78	0,60	0,75	0,72
Q56	0,62	0,63	0,63	0,66	0,66	0,66	0,53	0,54	0,73	0,68	0,69	0,63	0,70	0,73	0,70	0,74	0,64	0,76	0,69	0,79	0,80	1,00	0,81	0,66	0,75	0,75
Q57	0,55	0,55	0,57	0,62	0,62	0,59	0,49	0,51	0,67	0,68	0,63	0,62	0,70	0,69	0,63	0,67	0,60	0,71	0,70	0,78	0,78	0,81	1,00	0,65	0,70	0,73
Q58	0,47	0,46	0,51	0,68	0,62	0,44	0,41	0,40	0,72	0,62	0,79	0,68	0,71	0,70	0,54	0,60	0,88	0,73	0,72	0,59	0,60	0,66	0,65	1,00	0,77	0,80
Q59	0,73	0,60	0,79	0,80	0,81	0,63	0,53	0,53	0,83	0,76	0,80	0,74	0,75	0,81	0,77	0,85	0,77	0,80	0,70	0,76	0,75	0,75	0,70	0,77	1,00	0,80
Q60	0,55	0,51	0,58	0,73	0,66	0,59	0,60	0,53	0,79	0,71	0,75	0,69	0,74	0,72	0,59	0,67	0,81	0,83	0,86	0,73	0,72	0,75	0,73	0,80	0,80	1,00

Anexo V: Matrix de correlação Campus Pedra Branca

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q34	1,00	0,71	0,77	0,70	0,79	0,65	0,52	0,50	0,76	0,63	0,65	0,60	0,63	0,65	0,82	0,85	0,57	0,61	0,57	0,67	0,68	0,65	0,56	0,50	0,60	
Q36	0,71	1,00	0,65	0,57	0,58	0,49	0,39	0,34	0,57	0,45	0,56	0,54	0,56	0,63	0,68	0,65	0,47	0,50	0,43	0,56	0,53	0,55	0,47	0,46	0,49	
Q37	0,77	0,65	1,00	0,70	0,80	0,59	0,49	0,49	0,69	0,63	0,65	0,60	0,59	0,65	0,78	0,81	0,58	0,58	0,53	0,69	0,63	0,60	0,52	0,51	0,59	
Q38	0,70	0,57	0,70	1,00	0,81	0,55	0,47	0,47	0,84	0,65	0,74	0,72	0,66	0,73	0,67	0,78	0,71	0,69	0,65	0,64	0,66	0,62	0,56	0,63	0,69	
Q39	0,79	0,58	0,80	0,81	1,00	0,68	0,56	0,58	0,82	0,74	0,70	0,65	0,63	0,69	0,77	0,86	0,68	0,67	0,66	0,71	0,72	0,66	0,59	0,58	0,67	
Q40	0,65	0,49	0,59	0,55	0,68	1,00	0,68	0,75	0,63	0,68	0,55	0,46	0,52	0,50	0,64	0,64	0,48	0,57	0,49	0,58	0,60	0,60	0,53	0,42	0,59	
Q41	0,52	0,39	0,49	0,47	0,56	0,68	1,00	0,74	0,54	0,63	0,45	0,40	0,44	0,46	0,55	0,53	0,40	0,52	0,45	0,54	0,55	0,53	0,49	0,37	0,50	
Q42	0,50	0,34	0,49	0,47	0,58	0,75	0,74	1,00	0,55	0,70	0,47	0,39	0,44	0,42	0,50	0,53	0,39	0,54	0,49	0,52	0,56	0,53	0,49	0,37	0,51	
Q43	0,76	0,57	0,69	0,84	0,82	0,63	0,54	0,55	1,00	0,73	0,72	0,72	0,68	0,72	0,72	0,84	0,70	0,69	0,70	0,67	0,72	0,67	0,61	0,61	0,77	
Q44	0,63	0,45	0,63	0,65	0,74	0,68	0,63	0,70	0,73	1,00	0,67	0,56	0,58	0,58	0,64	0,72	0,61	0,72	0,68	0,64	0,63	0,63	0,60	0,58	0,71	
Q45	0,65	0,56	0,65	0,74	0,70	0,55	0,45	0,47	0,72	0,67	1,00	0,71	0,69	0,70	0,66	0,75	0,82	0,75	0,73	0,64	0,67	0,65	0,58	0,80	0,81	
Q46	0,60	0,54	0,60	0,72	0,65	0,46	0,40	0,39	0,72	0,56	0,71	1,00	0,78	0,82	0,64	0,70	0,69	0,61	0,63	0,56	0,60	0,58	0,51	0,63	0,73	
Q47	0,63	0,56	0,59	0,66	0,63	0,52	0,44	0,44	0,68	0,58	0,69	0,78	1,00	0,78	0,69	0,71	0,68	0,64	0,65	0,60	0,64	0,63	0,59	0,64	0,74	
Q48	0,65	0,63	0,65	0,73	0,69	0,50	0,46	0,42	0,72	0,58	0,70	0,82	0,78	1,00	0,74	0,75	0,67	0,65	0,61	0,63	0,64	0,62	0,56	0,63	0,76	
Q49	0,82	0,68	0,78	0,67	0,77	0,64	0,55	0,50	0,72	0,64	0,66	0,64	0,69	0,74	1,00	0,85	0,59	0,63	0,57	0,68	0,70	0,68	0,59	0,54	0,75	
Q50	0,85	0,65	0,81	0,78	0,86	0,64	0,53	0,53	0,84	0,72	0,75	0,70	0,71	0,75	0,85	1,00	0,71	0,72	0,68	0,72	0,74	0,72	0,62	0,63	0,85	
Q51	0,57	0,47	0,58	0,71	0,68	0,48	0,40	0,39	0,70	0,61	0,82	0,69	0,68	0,67	0,59	0,71	1,00	0,74	0,81	0,61	0,66	0,64	0,58	0,85	0,79	
Q52	0,61	0,50	0,58	0,69	0,67	0,57	0,52	0,54	0,69	0,72	0,75	0,61	0,64	0,65	0,63	0,72	0,74	1,00	0,78	0,67	0,68	0,71	0,65	0,70	0,81	
Q53	0,57	0,43	0,53	0,65	0,66	0,49	0,45	0,49	0,70	0,68	0,73	0,63	0,65	0,61	0,57	0,68	0,81	0,78	1,00	0,68	0,69	0,69	0,66	0,75	0,77	
Q54	0,67	0,56	0,69	0,64	0,71	0,58	0,54	0,52	0,67	0,64	0,64	0,56	0,60	0,63	0,68	0,72	0,61	0,67	0,68	1,00	0,77	0,74	0,73	0,57	0,74	
Q55	0,68	0,53	0,63	0,66	0,72	0,60	0,55	0,56	0,72	0,63	0,67	0,60	0,64	0,64	0,70	0,74	0,66	0,68	0,69	0,77	1,00	0,75	0,69	0,59	0,74	
Q56	0,65	0,55	0,60	0,62	0,66	0,60	0,53	0,53	0,67	0,63	0,65	0,58	0,63	0,62	0,68	0,72	0,64	0,71	0,69	0,74	0,75	1,00	0,84	0,64	0,74	
Q57	0,56	0,47	0,52	0,56	0,59	0,53	0,49	0,49	0,61	0,60	0,58	0,51	0,59	0,56	0,59	0,62	0,58	0,65	0,66	0,73	0,69	0,84	1,00	0,58	0,67	
Q58	0,50	0,46	0,51	0,63	0,58	0,42	0,37	0,37	0,61	0,58	0,80	0,63	0,64	0,63	0,54	0,63	0,85	0,70	0,75	0,57	0,59	0,64	0,58	1,00	0,77	
Q59	0,74	0,62	0,74	0,77	0,79	0,59	0,50	0,51	0,77	0,71	0,81	0,73	0,74	0,76	0,75	0,85	0,79	0,76	0,77	0,74	0,74	0,74	0,67	0,77	1,00	
Q60	0,60	0,49	0,59	0,69	0,67	0,54	0,49	0,51	0,71	0,68	0,76	0,64	0,67	0,65	0,63	0,71	0,79	0,81	0,85	0,67	0,71	0,73	0,67	0,77	0,80	

Anexo VI: Matrix de correlação (análise fatorial confirmatório) - Unisul Geral

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,59	0,55	0,45	0,38	0,40	0,35	0,27	0,30	0,44	0,37	0,42	0,33	0,39	0,42	0,51	0,51	0,31	0,40	0,31	0,42	0,42	0,47	0,33	0,37	0,45	0,37	
Q33	0,48	0,55	0,43	0,34	0,33	0,28	0,20	0,22	0,36	0,29	0,40	0,30	0,38	0,38	0,39	0,42	0,32	0,37	0,30	0,34	0,36	0,42	0,29	0,38	0,41	0,35	
Q34	1,00	0,71	0,77	0,67	0,79	0,60	0,44	0,44	0,74	0,62	0,64	0,57	0,57	0,66	0,80	0,84	0,56	0,61	0,52	0,64	0,67	0,59	0,48	0,49	0,72	0,54	
Q35	0,39	0,65	0,33	0,33	0,28	0,37	0,32	0,29	0,34	0,32	0,35	0,36	0,40	0,41	0,38	0,36	0,28	0,37	0,29	0,39	0,33	0,44	0,38	0,34	0,36	0,35	
Q36	0,71	1,00	0,63	0,55	0,49	0,34	0,34	0,32	0,56	0,47	0,57	0,50	0,53	0,60	0,68	0,65	0,47	0,52	0,43	0,57	0,55	0,56	0,46	0,47	0,63	0,48	
Q37	0,77	0,63	1,00	0,68	0,80	0,52	0,39	0,41	0,68	0,61	0,66	0,56	0,56	0,65	0,77	0,82	0,56	0,59	0,50	0,64	0,63	0,57	0,47	0,51	0,75	0,55	
Q38	0,67	0,55	0,68	1,00	0,76	0,48	0,42	0,41	0,84	0,64	0,70	0,69	0,61	0,70	0,65	0,77	0,70	0,67	0,63	0,58	0,64	0,59	0,51	0,60	0,72	0,67	
Q39	0,79	0,55	0,80	0,76	1,00	0,64	0,52	0,54	0,80	0,73	0,68	0,61	0,58	0,69	0,78	0,86	0,67	0,68	0,61	0,66	0,69	0,61	0,51	0,56	0,76	0,62	
Q40	0,60	0,49	0,52	0,48	0,64	1,00	0,65	0,73	0,61	0,69	0,50	0,40	0,48	0,52	0,63	0,62	0,43	0,55	0,50	0,61	0,63	0,59	0,51	0,40	0,55	0,51	
Q41	0,44	0,34	0,39	0,42	0,52	0,65	1,00	0,71	0,53	0,60	0,38	0,39	0,44	0,47	0,47	0,48	0,40	0,53	0,51	0,56	0,48	0,52	0,50	0,36	0,44	0,52	
Q42	0,44	0,32	0,41	0,41	0,54	0,73	0,71	1,00	0,54	0,69	0,42	0,36	0,39	0,43	0,48	0,49	0,33	0,54	0,50	0,52	0,54	0,51	0,46	0,33	0,45	0,51	
Q43	0,74	0,56	0,68	0,84	0,80	0,61	0,53	0,54	1,00	0,74	0,71	0,71	0,65	0,75	0,72	0,82	0,70	0,72	0,68	0,67	0,71	0,67	0,59	0,60	0,74	0,70	
Q44	0,62	0,47	0,61	0,64	0,73	0,69	0,60	0,69	0,74	1,00	0,67	0,54	0,56	0,61	0,66	0,72	0,58	0,73	0,66	0,62	0,66	0,63	0,59	0,55	0,69	0,68	
Q45	0,64	0,57	0,66	0,70	0,68	0,50	0,38	0,42	0,71	0,67	1,00	0,67	0,66	0,69	0,67	0,76	0,80	0,73	0,66	0,57	0,66	0,62	0,50	0,77	0,79	0,72	
Q46	0,57	0,50	0,56	0,69	0,61	0,40	0,39	0,36	0,71	0,54	0,67	1,00	0,77	0,83	0,58	0,67	0,69	0,64	0,62	0,52	0,56	0,56	0,51	0,63	0,71	0,66	
Q47	0,57	0,53	0,56	0,61	0,58	0,48	0,44	0,39	0,65	0,56	0,66	0,77	1,00	0,79	0,61	0,65	0,67	0,61	0,60	0,56	0,59	0,61	0,55	0,61	0,68	0,65	
Q48	0,66	0,60	0,65	0,70	0,69	0,52	0,47	0,43	0,75	0,61	0,69	0,83	0,79	1,00	0,73	0,76	0,69	0,68	0,62	0,64	0,65	0,65	0,59	0,65	0,75	0,67	
Q49	0,80	0,68	0,77	0,65	0,78	0,63	0,47	0,48	0,72	0,66	0,67	0,58	0,61	0,73	1,00	0,86	0,57	0,65	0,55	0,68	0,72	0,64	0,54	0,52	0,77	0,58	
Q50	0,84	0,65	0,82	0,77	0,86	0,62	0,48	0,49	0,82	0,72	0,76	0,67	0,65	0,76	0,86	1,00	0,69	0,73	0,64	0,69	0,74	0,68	0,57	0,61	0,83	0,68	
Q51	0,56	0,47	0,56	0,70	0,67	0,43	0,40	0,33	0,70	0,58	0,80	0,69	0,67	0,69	0,57	0,69	1,00	0,74	0,75	0,56	0,61	0,61	0,53	0,83	0,77	0,75	
Q52	0,61	0,52	0,59	0,67	0,68	0,55	0,53	0,54	0,72	0,73	0,73	0,64	0,61	0,68	0,65	0,73	0,74	1,00	0,78	0,66	0,70	0,72	0,62	0,71	0,75	0,80	
Q53	0,52	0,43	0,50	0,63	0,61	0,50	0,51	0,50	0,68	0,66	0,66	0,62	0,60	0,62	0,55	0,64	0,75	0,78	1,00	0,67	0,66	0,69	0,66	0,70	0,71	0,83	
Q54	0,64	0,57	0,64	0,58	0,66	0,61	0,56	0,52	0,67	0,62	0,57	0,52	0,56	0,64	0,68	0,69	0,56	0,66	0,67	1,00	0,78	0,74	0,74	0,51	0,70	0,64	
Q55	0,67	0,55	0,63	0,64	0,69	0,63	0,48	0,54	0,71	0,66	0,66	0,56	0,59	0,65	0,72	0,74	0,61	0,70	0,66	0,78	1,00	0,72	0,66	0,56	0,73	0,66	
Q56	0,59	0,56	0,57	0,59	0,61	0,59	0,52	0,51	0,67	0,63	0,62	0,56	0,61	0,65	0,64	0,68	0,61	0,72	0,69	0,74	0,72	1,00	0,82	0,63	0,72	0,74	
Q57	0,48	0,46	0,47	0,51	0,51	0,51	0,50	0,46	0,59	0,59	0,50	0,51	0,55	0,59	0,54	0,57	0,53	0,62	0,66	0,74	0,66	0,82	1,00	0,55	0,62	0,65	
Q58	0,49	0,47	0,51	0,60	0,56	0,40	0,36	0,33	0,60	0,55	0,77	0,63	0,61	0,65	0,52	0,61	0,83	0,71	0,70	0,51	0,56	0,63	0,55	1,00	0,75	0,77	
Q59	0,72	0,63	0,75	0,72	0,76	0,55	0,44	0,45	0,74	0,69	0,79	0,71	0,68	0,75	0,77	0,83	0,77	0,75	0,71	0,70	0,73	0,72	0,62	0,62	0,75	1,00	0,79
Q60	0,54	0,48	0,55	0,67	0,62	0,51	0,52	0,51	0,70	0,68	0,72	0,66	0,65	0,67	0,58	0,68	0,75	0,80	0,83	0,64	0,66	0,74	0,65	0,77	0,79	1,00	

NO. CASES. 675

Anexo VII: Matrix de correlação (análise fatorial confirmatório) – Campus Tubarão

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q32	0,54	0,51	0,44	0,37	0,37	0,31	0,27	0,24	0,39	0,39	0,42	0,34	0,39	0,46	0,50	0,48	0,30	0,41	0,31	0,39	0,37	0,44	0,32	0,34	0,36
Q33	0,43	0,51	0,39	0,30	0,29	0,26	0,22	0,16	0,30	0,30	0,37	0,33	0,41	0,43	0,41	0,40	0,33	0,36	0,30	0,34	0,33	0,43	0,31	0,38	0,40
Q34	1,00	0,67	0,78	0,70	0,80	0,59	0,36	0,41	0,72	0,61	0,65	0,52	0,54	0,65	0,82	0,86	0,55	0,58	0,50	0,62	0,66	0,54	0,46	0,43	0,48
Q35	0,33	0,63	0,29	0,28	0,25	0,40	0,37	0,30	0,29	0,32	0,27	0,29	0,40	0,37	0,34	0,31	0,25	0,34	0,30	0,39	0,29	0,40	0,36	0,26	0,29
Q36	0,67	1,00	0,61	0,53	0,55	0,52	0,33	0,32	0,56	0,50	0,54	0,47	0,56	0,60	0,69	0,64	0,48	0,52	0,46	0,59	0,54	0,56	0,48	0,43	0,48
Q37	0,78	0,61	1,00	0,70	0,82	0,53	0,36	0,39	0,67	0,62	0,66	0,52	0,55	0,63	0,79	0,82	0,56	0,59	0,49	0,62	0,63	0,58	0,47	0,45	0,48
Q38	0,70	0,53	0,70	1,00	0,76	0,46	0,34	0,39	0,83	0,66	0,72	0,68	0,58	0,69	0,69	0,79	0,69	0,66	0,63	0,57	0,63	0,56	0,47	0,56	0,63
Q39	0,80	0,55	0,82	0,76	1,00	0,63	0,41	0,50	0,78	0,70	0,67	0,58	0,54	0,64	0,79	0,87	0,64	0,66	0,57	0,65	0,71	0,57	0,48	0,48	0,54
Q40	0,59	0,52	0,53	0,46	0,63	1,00	0,61	0,72	0,58	0,63	0,47	0,37	0,44	0,50	0,63	0,62	0,37	0,51	0,46	0,60	0,64	0,53	0,49	0,28	0,43
Q41	0,36	0,33	0,36	0,34	0,41	0,61	1,00	0,72	0,47	0,55	0,33	0,30	0,41	0,43	0,40	0,41	0,30	0,51	0,48	0,57	0,45	0,52	0,54	0,28	0,39
Q42	0,41	0,32	0,39	0,39	0,50	0,72	0,72	1,00	0,53	0,70	0,38	0,28	0,33	0,38	0,45	0,49	0,29	0,53	0,47	0,51	0,51	0,47	0,46	0,23	0,42
Q43	0,72	0,56	0,67	0,83	0,78	0,58	0,47	0,53	1,00	0,75	0,73	0,71	0,63	0,74	0,72	0,83	0,69	0,72	0,69	0,64	0,69	0,60	0,55	0,56	0,77
Q44	0,61	0,50	0,62	0,66	0,70	0,63	0,55	0,70	0,75	1,00	0,70	0,53	0,54	0,61	0,65	0,73	0,60	0,74	0,67	0,63	0,65	0,61	0,58	0,50	0,70
Q45	0,65	0,54	0,66	0,72	0,67	0,47	0,33	0,38	0,73	0,70	1,00	0,64	0,62	0,67	0,65	0,74	0,77	0,70	0,64	0,56	0,60	0,59	0,48	0,71	0,77
Q46	0,52	0,47	0,52	0,68	0,58	0,37	0,30	0,28	0,71	0,53	0,64	1,00	0,73	0,82	0,54	0,64	0,68	0,62	0,62	0,49	0,49	0,53	0,46	0,61	0,67
Q47	0,54	0,56	0,55	0,58	0,54	0,44	0,41	0,33	0,63	0,54	0,62	0,73	1,00	0,79	0,56	0,61	0,64	0,58	0,58	0,56	0,51	0,57	0,51	0,59	0,64
Q48	0,65	0,60	0,63	0,69	0,64	0,50	0,43	0,38	0,74	0,61	0,67	0,82	0,79	1,00	0,68	0,73	0,68	0,66	0,64	0,63	0,59	0,63	0,57	0,62	0,72
Q49	0,82	0,69	0,79	0,69	0,79	0,63	0,40	0,45	0,72	0,65	0,65	0,54	0,56	0,68	1,00	0,88	0,54	0,63	0,54	0,69	0,71	0,62	0,53	0,43	0,76
Q50	0,86	0,64	0,82	0,79	0,87	0,62	0,41	0,49	0,83	0,73	0,74	0,64	0,61	0,73	0,88	1,00	0,68	0,71	0,62	0,70	0,74	0,65	0,55	0,54	0,84
Q51	0,55	0,48	0,56	0,69	0,64	0,37	0,30	0,29	0,69	0,60	0,77	0,68	0,64	0,68	0,54	0,68	1,00	0,70	0,73	0,55	0,54	0,57	0,51	0,81	0,76
Q52	0,58	0,52	0,59	0,66	0,66	0,51	0,51	0,53	0,72	0,74	0,70	0,62	0,58	0,66	0,63	0,71	0,70	1,00	0,78	0,66	0,66	0,69	0,59	0,66	0,74
Q53	0,50	0,46	0,49	0,63	0,57	0,46	0,48	0,47	0,69	0,67	0,64	0,62	0,58	0,64	0,54	0,62	0,73	0,78	1,00	0,66	0,60	0,67	0,66	0,65	0,68
Q54	0,62	0,59	0,62	0,57	0,65	0,60	0,57	0,51	0,64	0,63	0,56	0,49	0,56	0,63	0,69	0,70	0,55	0,66	0,66	1,00	0,77	0,78	0,79	0,47	0,68
Q55	0,66	0,54	0,63	0,63	0,71	0,64	0,45	0,51	0,69	0,65	0,60	0,49	0,51	0,59	0,71	0,74	0,54	0,66	0,60	0,77	1,00	0,72	0,63	0,44	0,70
Q56	0,54	0,56	0,58	0,56	0,57	0,53	0,52	0,47	0,60	0,61	0,59	0,53	0,57	0,63	0,62	0,65	0,57	0,69	0,67	0,78	0,72	1,00	0,82	0,58	0,71
Q57	0,46	0,48	0,47	0,47	0,48	0,49	0,54	0,46	0,55	0,58	0,48	0,46	0,51	0,57	0,53	0,55	0,51	0,59	0,66	0,79	0,63	0,82	1,00	0,48	0,59
Q58	0,43	0,43	0,45	0,56	0,48	0,28	0,28	0,23	0,56	0,50	0,71	0,61	0,59	0,62	0,43	0,54	0,81	0,66	0,65	0,47	0,44	0,58	0,48	1,00	0,70
Q59	0,73	0,63	0,75	0,74	0,77	0,53	0,39	0,42	0,77	0,70	0,77	0,67	0,64	0,72	0,76	0,84	0,76	0,74	0,68	0,68	0,70	0,71	0,59	0,70	1,00
Q60	0,48	0,48	0,48	0,63	0,54	0,43	0,46	0,47	0,69	0,68	0,66	0,63	0,58	0,64	0,53	0,63	0,71	0,78	0,83	0,60	0,59	0,71	0,65	0,70	1,00

NO. CASES.

675

Anexo VIII: Matrix de correlação (análise fatorial confirmatório) – Campus Araranguá

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,52	0,64	0,37	0,27	0,34	0,28	0,21	0,27	0,32	0,23	0,34	0,24	0,30	0,34	0,42	0,41	0,27	0,38	0,30	0,36	0,32	0,44	0,29	0,31	0,37	0,32
Q33	0,41	0,54	0,33	0,30	0,27	0,36	0,23	0,28	0,34	0,26	0,42	0,31	0,39	0,36	0,39	0,34	0,32	0,44	0,38	0,35	0,38	0,48	0,35	0,39	0,38	0,42
Q34	1,00	0,72	0,79	0,67	0,77	0,56	0,44	0,44	0,70	0,58	0,61	0,55	0,56	0,65	0,81	0,82	0,55	0,58	0,55	0,68	0,64	0,63	0,57	0,49	0,73	0,56
Q35	0,36	0,67	0,29	0,29	0,25	0,39	0,37	0,39	0,35	0,34	0,36	0,33	0,41	0,34	0,36	0,34	0,30	0,43	0,41	0,41	0,40	0,53	0,45	0,39	0,38	0,45
Q36	0,72	1,00	0,61	0,50	0,56	0,48	0,41	0,38	0,53	0,45	0,54	0,48	0,56	0,58	0,66	0,64	0,48	0,54	0,46	0,59	0,56	0,65	0,57	0,47	0,61	0,52
Q37	0,79	0,61	1,00	0,70	0,85	0,57	0,37	0,44	0,69	0,64	0,67	0,57	0,61	0,66	0,81	0,85	0,59	0,63	0,55	0,71	0,64	0,63	0,57	0,52	0,79	0,58
Q38	0,67	0,50	0,70	1,00	0,77	0,54	0,43	0,45	0,87	0,67	0,76	0,73	0,65	0,75	0,65	0,74	0,75	0,72	0,70	0,68	0,64	0,64	0,62	0,68	0,78	0,73
Q39	0,77	0,56	0,85	0,77	1,00	0,63	0,48	0,51	0,78	0,71	0,69	0,63	0,64	0,70	0,77	0,86	0,70	0,64	0,73	0,66	0,67	0,62	0,62	0,63	0,80	0,67
Q40	0,56	0,48	0,57	0,54	0,63	1,00	0,66	0,77	0,62	0,70	0,53	0,48	0,51	0,55	0,62	0,60	0,47	0,59	0,54	0,66	0,65	0,66	0,59	0,44	0,62	0,58
Q41	0,44	0,41	0,37	0,43	0,48	0,66	1,00	0,66	0,57	0,58	0,44	0,42	0,44	0,43	0,46	0,46	0,46	0,56	0,56	0,55	0,56	0,53	0,49	0,42	0,53	0,59
Q42	0,44	0,38	0,44	0,45	0,51	0,77	0,66	1,00	0,56	0,66	0,44	0,44	0,42	0,44	0,46	0,48	0,39	0,55	0,50	0,53	0,53	0,50	0,41	0,52	0,51	
Q43	0,70	0,53	0,69	0,87	0,78	0,62	0,57	0,56	1,00	0,75	0,76	0,76	0,70	0,79	0,71	0,78	0,76	0,75	0,73	0,71	0,69	0,71	0,65	0,71	0,81	0,78
Q44	0,58	0,45	0,64	0,67	0,71	0,70	0,58	0,66	0,75	1,00	0,69	0,60	0,64	0,68	0,68	0,71	0,63	0,72	0,66	0,68	0,66	0,68	0,66	0,62	0,73	0,70
Q45	0,61	0,54	0,67	0,76	0,69	0,53	0,44	0,44	0,76	0,69	1,00	0,71	0,71	0,74	0,68	0,72	0,81	0,77	0,70	0,67	0,65	0,70	0,64	0,79	0,80	0,76
Q46	0,55	0,48	0,57	0,73	0,63	0,48	0,42	0,44	0,76	0,60	0,71	1,00	0,81	0,84	0,58	0,65	0,66	0,68	0,63	0,62	0,63	0,62	0,61	0,68	0,74	0,69
Q47	0,56	0,56	0,61	0,65	0,64	0,51	0,44	0,42	0,70	0,64	0,71	0,81	1,00	0,81	0,66	0,65	0,67	0,72	0,64	0,68	0,67	0,70	0,69	0,71	0,75	0,75
Q48	0,65	0,58	0,66	0,75	0,70	0,55	0,43	0,44	0,79	0,68	0,74	0,84	0,81	1,00	0,72	0,77	0,69	0,70	0,61	0,69	0,71	0,71	0,68	0,70	0,79	0,71
Q49	0,81	0,66	0,81	0,65	0,77	0,62	0,46	0,46	0,71	0,68	0,68	0,58	0,66	0,72	1,00	0,85	0,57	0,65	0,55	0,74	0,67	0,68	0,63	0,54	0,77	0,59
Q50	0,82	0,64	0,85	0,74	0,86	0,60	0,46	0,48	0,78	0,71	0,72	0,65	0,65	0,77	0,85	1,00	0,67	0,68	0,62	0,77	0,70	0,73	0,66	0,61	0,84	0,66
Q51	0,55	0,48	0,59	0,75	0,70	0,47	0,46	0,39	0,76	0,63	0,81	0,66	0,67	0,69	0,57	0,67	1,00	0,75	0,78	0,62	0,58	0,64	0,60	0,88	0,77	0,81
Q52	0,58	0,54	0,63	0,72	0,70	0,59	0,56	0,55	0,75	0,72	0,77	0,68	0,72	0,70	0,65	0,68	0,75	1,00	0,79	0,71	0,72	0,75	0,70	0,74	0,80	0,83
Q53	0,55	0,46	0,55	0,70	0,64	0,54	0,56	0,50	0,73	0,66	0,70	0,63	0,64	0,61	0,55	0,62	0,78	0,79	1,00	0,69	0,68	0,69	0,71	0,73	0,72	0,86
Q54	0,68	0,59	0,71	0,68	0,73	0,66	0,55	0,53	0,71	0,68	0,67	0,62	0,68	0,69	0,74	0,77	0,62	0,71	0,69	1,00	0,84	0,79	0,79	0,60	0,77	0,73
Q55	0,64	0,56	0,64	0,64	0,66	0,65	0,56	0,53	0,69	0,66	0,65	0,63	0,67	0,71	0,67	0,70	0,58	0,72	0,68	0,84	1,00	0,78	0,78	0,59	0,75	0,72
Q56	0,63	0,65	0,63	0,64	0,67	0,66	0,53	0,53	0,71	0,68	0,70	0,62	0,70	0,71	0,68	0,73	0,64	0,75	0,69	0,79	0,78	1,00	0,81	0,67	0,75	0,75
Q57	0,57	0,57	0,57	0,62	0,62	0,59	0,49	0,50	0,65	0,66	0,64	0,61	0,69	0,68	0,63	0,66	0,60	0,70	0,71	0,79	0,78	0,81	1,00	0,65	0,71	0,72
Q58	0,49	0,47	0,52	0,68	0,63	0,44	0,42	0,41	0,71	0,62	0,79	0,68	0,71	0,70	0,54	0,61	0,88	0,74	0,73	0,60	0,59	0,67	0,65	1,00	0,77	0,79
Q59	0,73	0,61	0,79	0,78	0,80	0,62	0,53	0,52	0,81	0,73	0,80	0,74	0,75	0,79	0,77	0,84	0,77	0,80	0,72	0,77	0,75	0,75	0,71	0,77	1,00	0,80
Q60	0,56	0,52	0,58	0,73	0,67	0,58	0,59	0,51	0,78	0,70	0,76	0,69	0,75	0,71	0,59	0,66	0,81	0,83	0,86	0,73	0,72	0,75	0,72	0,79	0,80	1,00

NO. CASES.

249

Anexo IX: Matrix de correlação (análise fatorial confirmatório) – Campus Pedra Branca

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,59	0,52	0,44	0,34	0,40	0,36	0,32	0,30	0,43	0,37	0,37	0,35	0,37	0,38	0,48	0,50	0,28	0,35	0,30	0,42	0,41	0,47	0,36	0,32	0,43	0,34
Q33	0,44	0,49	0,38	0,30	0,32	0,27	0,20	0,21	0,34	0,25	0,33	0,31	0,32	0,34	0,37	0,38	0,26	0,32	0,25	0,33	0,32	0,39	0,31	0,31	0,35	0,32
Q34	1,00	0,73	0,78	0,70	0,79	0,64	0,52	0,48	0,78	0,64	0,65	0,62	0,62	0,66	0,83	0,86	0,57	0,60	0,58	0,67	0,68	0,66	0,56	0,51	0,75	0,60
Q35	0,43	0,66	0,38	0,36	0,32	0,31	0,22	0,22	0,36	0,28	0,37	0,36	0,37	0,42	0,38	0,38	0,28	0,34	0,26	0,38	0,34	0,40	0,34	0,32	0,38	0,33
Q36	0,73	1,00	0,66	0,56	0,58	0,49	0,38	0,33	0,58	0,45	0,55	0,55	0,55	0,64	0,68	0,66	0,46	0,48	0,43	0,55	0,52	0,55	0,46	0,46	0,63	0,47
Q37	0,78	0,66	1,00	0,70	0,80	0,57	0,48	0,49	0,70	0,64	0,66	0,61	0,59	0,66	0,78	0,82	0,58	0,58	0,54	0,70	0,64	0,61	0,52	0,53	0,75	0,59
Q38	0,70	0,56	0,70	1,00	0,81	0,55	0,47	0,46	0,84	0,66	0,74	0,72	0,67	0,73	0,67	0,77	0,72	0,69	0,66	0,65	0,66	0,62	0,57	0,64	0,78	0,70
Q39	0,79	0,58	0,80	0,81	1,00	0,69	0,56	0,58	0,83	0,76	0,72	0,66	0,63	0,70	0,77	0,86	0,68	0,68	0,72	0,72	0,72	0,68	0,60	0,59	0,80	0,68
Q40	0,64	0,49	0,57	0,55	0,69	1,00	0,68	0,75	0,65	0,69	0,56	0,48	0,52	0,51	0,65	0,65	0,49	0,59	0,52	0,60	0,63	0,62	0,55	0,45	0,60	0,56
Q41	0,52	0,38	0,48	0,47	0,56	0,68	1,00	0,74	0,55	0,64	0,45	0,43	0,46	0,47	0,55	0,53	0,41	0,52	0,46	0,54	0,56	0,54	0,49	0,40	0,50	0,50
Q42	0,48	0,33	0,49	0,46	0,58	0,75	0,74	1,00	0,55	0,69	0,47	0,41	0,45	0,43	0,50	0,52	0,40	0,55	0,50	0,53	0,56	0,53	0,50	0,39	0,51	0,53
Q43	0,78	0,58	0,70	0,84	0,83	0,65	0,55	0,55	1,00	0,74	0,73	0,74	0,68	0,74	0,72	0,84	0,71	0,70	0,71	0,68	0,73	0,69	0,63	0,62	0,78	0,72
Q44	0,64	0,45	0,64	0,66	0,76	0,69	0,64	0,69	0,74	1,00	0,68	0,57	0,60	0,59	0,64	0,73	0,63	0,73	0,70	0,65	0,65	0,65	0,61	0,60	0,71	0,70
Q45	0,65	0,55	0,66	0,74	0,72	0,56	0,45	0,47	0,73	0,68	1,00	0,71	0,70	0,70	0,67	0,75	0,83	0,76	0,75	0,65	0,69	0,66	0,58	0,80	0,82	0,77
Q46	0,62	0,55	0,61	0,72	0,66	0,48	0,43	0,41	0,74	0,57	0,71	1,00	0,79	0,82	0,65	0,70	0,69	0,62	0,65	0,57	0,62	0,59	0,51	0,63	0,73	0,65
Q47	0,62	0,55	0,59	0,67	0,63	0,52	0,46	0,45	0,68	0,60	0,70	0,79	1,00	0,78	0,67	0,70	0,68	0,64	0,66	0,60	0,65	0,63	0,60	0,63	0,73	0,67
Q48	0,66	0,64	0,66	0,73	0,70	0,51	0,47	0,43	0,74	0,59	0,70	0,82	0,78	1,00	0,74	0,75	0,67	0,65	0,62	0,64	0,66	0,63	0,57	0,64	0,76	0,66
Q49	0,83	0,68	0,78	0,67	0,77	0,65	0,55	0,50	0,72	0,64	0,67	0,65	0,67	0,74	1,00	0,85	0,58	0,63	0,57	0,68	0,70	0,68	0,59	0,54	0,76	0,61
Q50	0,86	0,66	0,82	0,77	0,86	0,65	0,53	0,52	0,84	0,73	0,75	0,70	0,70	0,75	0,85	1,00	0,70	0,72	0,67	0,73	0,74	0,72	0,62	0,63	0,85	0,71
Q51	0,57	0,46	0,58	0,72	0,68	0,49	0,41	0,40	0,71	0,63	0,83	0,69	0,68	0,67	0,58	0,70	1,00	0,74	0,81	0,63	0,66	0,64	0,57	0,86	0,78	0,78
Q52	0,60	0,48	0,58	0,69	0,68	0,59	0,52	0,55	0,70	0,73	0,76	0,62	0,64	0,65	0,63	0,72	0,74	1,00	0,78	0,67	0,67	0,71	0,65	0,70	0,75	0,81
Q53	0,58	0,43	0,54	0,66	0,68	0,52	0,46	0,50	0,71	0,70	0,75	0,65	0,66	0,62	0,57	0,67	0,81	0,78	1,00	0,69	0,68	0,68	0,65	0,76	0,77	0,85
Q54	0,67	0,55	0,70	0,65	0,72	0,60	0,54	0,53	0,68	0,65	0,65	0,57	0,60	0,64	0,68	0,73	0,63	0,67	0,69	1,00	0,78	0,73	0,72	0,57	0,74	0,68
Q55	0,68	0,52	0,64	0,66	0,72	0,63	0,56	0,56	0,73	0,65	0,69	0,62	0,65	0,66	0,70	0,74	0,66	0,67	0,68	0,78	1,00	0,74	0,69	0,59	0,74	0,70
Q56	0,66	0,55	0,61	0,62	0,68	0,62	0,54	0,53	0,69	0,65	0,66	0,59	0,63	0,68	0,72	0,64	0,64	0,71	0,68	0,73	0,74	1,00	0,83	0,63	0,73	0,72
Q57	0,56	0,46	0,52	0,57	0,60	0,55	0,49	0,50	0,63	0,61	0,58	0,51	0,60	0,57	0,59	0,62	0,57	0,65	0,65	0,72	0,69	0,83	1,00	0,57	0,65	0,67
Q58	0,51	0,46	0,53	0,64	0,59	0,45	0,40	0,39	0,62	0,60	0,80	0,63	0,63	0,64	0,54	0,63	0,86	0,70	0,76	0,57	0,59	0,63	0,57	1,00	0,77	0,78
Q59	0,75	0,63	0,75	0,78	0,80	0,60	0,50	0,51	0,78	0,71	0,82	0,73	0,73	0,76	0,76	0,85	0,78	0,75	0,77	0,74	0,74	0,73	0,65	0,77	1,00	0,80
Q60	0,60	0,47	0,59	0,70	0,68	0,56	0,50	0,53	0,72	0,70	0,77	0,65	0,67	0,66	0,61	0,71	0,78	0,81	0,85	0,68	0,70	0,72	0,67	0,78	0,80	1,00

NO. CASES. 675

Anexo X: Matrix de correlação (análise fatorial confirmatório) – Unisul Geral

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,56	0,58	0,41	0,30	0,39	0,33	0,27	0,33	0,39	0,34	0,39	0,27	0,35	0,33	0,45	0,45	0,31	0,36	0,29	0,37	0,39	0,44	0,32	0,35	0,38	0,34
Q33	0,42	0,57	0,39	0,29	0,29	0,33	0,28	0,28	0,36	0,31	0,42	0,28	0,38	0,35	0,37	0,34	0,38	0,29	0,37	0,35	0,46	0,36	0,36	0,38	0,37	0,36
Q34	1,00	0,68	0,73	0,68	0,78	0,53	0,44	0,44	0,70	0,59	0,62	0,56	0,53	0,61	0,78	0,82	0,55	0,54	0,62	0,62	0,60	0,53	0,50	0,70	0,51	
Q35	0,42	0,66	0,37	0,31	0,31	0,42	0,38	0,39	0,35	0,34	0,38	0,36	0,40	0,42	0,37	0,35	0,30	0,36	0,31	0,44	0,34	0,44	0,40	0,36	0,33	0,37
Q36	0,68	1,00	0,62	0,49	0,54	0,50	0,42	0,36	0,54	0,44	0,55	0,47	0,53	0,59	0,64	0,61	0,48	0,48	0,41	0,57	0,52	0,59	0,53	0,47	0,59	0,47
Q37	0,73	0,62	1,00	0,68	0,81	0,51	0,41	0,44	0,65	0,59	0,65	0,51	0,50	0,58	0,75	0,79	0,56	0,54	0,44	0,59	0,59	0,54	0,48	0,49	0,71	0,51
Q38	0,68	0,49	0,68	1,00	0,75	0,45	0,39	0,44	0,81	0,62	0,70	0,65	0,53	0,65	0,64	0,76	0,67	0,62	0,58	0,58	0,61	0,55	0,52	0,57	0,72	0,61
Q39	0,78	0,54	0,81	0,75	1,00	0,62	0,49	0,54	0,75	0,69	0,66	0,58	0,52	0,65	0,75	0,84	0,62	0,64	0,56	0,64	0,68	0,58	0,54	0,52	0,76	0,57
Q40	0,53	0,50	0,51	0,45	0,62	1,00	0,60	0,72	0,54	0,60	0,47	0,34	0,37	0,44	0,57	0,56	0,35	0,50	0,38	0,51	0,53	0,53	0,45	0,33	0,47	0,41
Q41	0,44	0,42	0,41	0,39	0,49	0,60	1,00	0,71	0,48	0,52	0,38	0,42	0,43	0,47	0,49	0,46	0,35	0,51	0,48	0,56	0,49	0,53	0,53	0,37	0,45	0,49
Q42	0,44	0,36	0,44	0,44	0,54	0,72	0,71	1,00	0,55	0,69	0,40	0,39	0,40	0,45	0,47	0,49	0,32	0,53	0,46	0,56	0,58	0,51	0,48	0,33	0,43	0,48
Q43	0,70	0,54	0,65	0,81	0,75	0,54	0,48	0,55	1,00	0,74	0,71	0,70	0,61	0,72	0,68	0,82	0,70	0,70	0,68	0,65	0,71	0,67	0,60	0,61	0,75	0,67
Q44	0,59	0,44	0,59	0,62	0,69	0,60	0,52	0,69	0,74	1,00	0,65	0,56	0,55	0,60	0,61	0,71	0,60	0,71	0,63	0,61	0,67	0,62	0,60	0,55	0,68	0,65
Q45	0,62	0,55	0,65	0,70	0,66	0,47	0,38	0,40	0,71	0,65	1,00	0,70	0,66	0,70	0,67	0,72	0,82	0,71	0,65	0,57	0,63	0,62	0,54	0,79	0,77	0,71
Q46	0,56	0,47	0,51	0,65	0,58	0,34	0,42	0,39	0,70	0,56	0,70	1,00	0,78	0,81	0,56	0,63	0,68	0,61	0,62	0,52	0,54	0,57	0,57	0,64	0,68	0,63
Q47	0,53	0,53	0,50	0,53	0,52	0,37	0,43	0,40	0,61	0,55	0,66	0,78	1,00	0,78	0,60	0,61	0,64	0,61	0,58	0,55	0,57	0,59	0,60	0,63	0,63	0,62
Q48	0,61	0,59	0,58	0,65	0,65	0,44	0,47	0,45	0,72	0,60	0,70	0,81	0,78	1,00	0,72	0,71	0,65	0,63	0,58	0,60	0,61	0,60	0,62	0,64	0,72	0,63
Q49	0,78	0,64	0,75	0,64	0,75	0,57	0,49	0,47	0,68	0,61	0,67	0,56	0,60	0,72	1,00	0,83	0,56	0,61	0,52	0,63	0,65	0,62	0,60	0,52	0,72	0,56
Q50	0,82	0,61	0,79	0,76	0,84	0,56	0,46	0,49	0,82	0,71	0,72	0,63	0,61	0,71	0,83	1,00	0,69	0,68	0,62	0,67	0,71	0,67	0,58	0,61	0,83	0,64
Q51	0,55	0,48	0,56	0,67	0,62	0,35	0,35	0,32	0,70	0,60	0,82	0,68	0,64	0,65	0,56	0,69	1,00	0,74	0,77	0,57	0,62	0,64	0,57	0,86	0,78	0,76
Q52	0,54	0,48	0,54	0,62	0,64	0,50	0,51	0,53	0,70	0,71	0,71	0,61	0,61	0,63	0,61	0,68	0,74	1,00	0,76	0,60	0,65	0,71	0,63	0,71	0,76	0,76
Q53	0,49	0,41	0,44	0,58	0,56	0,38	0,48	0,46	0,68	0,63	0,65	0,62	0,58	0,58	0,52	0,62	0,77	0,76	1,00	0,66	0,66	0,69	0,68	0,72	0,69	0,81
Q54	0,62	0,57	0,59	0,58	0,64	0,51	0,56	0,56	0,65	0,61	0,57	0,52	0,55	0,60	0,63	0,67	0,57	0,60	0,66	1,00	0,77	0,74	0,78	0,53	0,66	0,63
Q55	0,62	0,52	0,59	0,61	0,68	0,53	0,49	0,58	0,71	0,67	0,63	0,54	0,57	0,61	0,65	0,71	0,62	0,65	0,66	0,77	1,00	0,73	0,69	0,56	0,69	0,64
Q56	0,60	0,59	0,54	0,55	0,58	0,53	0,53	0,51	0,67	0,62	0,62	0,57	0,59	0,60	0,62	0,67	0,64	0,71	0,69	0,74	0,73	1,00	0,83	0,67	0,76	0,72
Q57	0,53	0,53	0,48	0,52	0,54	0,45	0,53	0,48	0,60	0,60	0,54	0,57	0,60	0,62	0,60	0,58	0,57	0,63	0,68	0,78	0,69	0,83	1,00	0,60	0,65	0,68
Q58	0,50	0,47	0,49	0,57	0,52	0,33	0,37	0,33	0,61	0,55	0,79	0,64	0,63	0,64	0,52	0,61	0,86	0,71	0,72	0,53	0,56	0,67	0,60	1,00	0,77	0,80
Q59	0,70	0,59	0,71	0,72	0,76	0,47	0,45	0,43	0,75	0,68	0,77	0,68	0,63	0,72	0,72	0,83	0,78	0,76	0,69	0,66	0,69	0,76	0,65	0,77	1,00	0,74
Q60	0,51	0,47	0,51	0,61	0,57	0,41	0,49	0,48	0,67	0,65	0,71	0,63	0,62	0,63	0,56	0,64	0,76	0,76	0,81	0,63	0,64	0,72	0,68	0,80	0,74	1,00

NO. CASES. 300

Anexo XI: Matrix de correlação (análise fatorial confirmatório) – Campus Tubarão

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,53	0,49	0,39	0,37	0,33	0,30	0,25	0,22	0,38	0,36	0,41	0,37	0,40	0,46	0,43	0,42	0,35	0,37	0,37	0,34	0,31	0,40	0,34	0,39	0,42	0,37
Q33	0,48	0,49	0,45	0,40	0,33	0,28	0,24	0,20	0,35	0,34	0,44	0,39	0,46	0,47	0,43	0,41	0,40	0,41	0,36	0,35	0,35	0,44	0,37	0,46	0,46	0,38
Q34	1,00	0,69	0,79	0,73	0,81	0,62	0,41	0,44	0,74	0,63	0,70	0,58	0,58	0,68	0,82	0,86	0,60	0,64	0,57	0,65	0,69	0,60	0,53	0,51	0,75	0,53
Q35	0,34	0,57	0,33	0,33	0,26	0,35	0,39	0,27	0,30	0,33	0,31	0,33	0,37	0,38	0,36	0,31	0,27	0,31	0,31	0,41	0,31	0,44	0,45	0,31	0,30	0,32
Q36	0,69	1,00	0,63	0,57	0,58	0,54	0,36	0,34	0,59	0,55	0,59	0,48	0,53	0,61	0,73	0,67	0,52	0,57	0,52	0,63	0,59	0,61	0,57	0,46	0,64	0,51
Q37	0,79	0,63	1,00	0,73	0,83	0,55	0,38	0,38	0,67	0,60	0,72	0,57	0,62	0,65	0,77	0,83	0,63	0,64	0,54	0,64	0,64	0,62	0,53	0,57	0,79	0,55
Q38	0,73	0,57	0,73	1,00	0,76	0,51	0,46	0,44	0,81	0,67	0,74	0,68	0,62	0,69	0,70	0,81	0,71	0,70	0,67	0,61	0,66	0,63	0,52	0,59	0,77	0,66
Q39	0,81	0,58	0,83	0,76	1,00	0,67	0,51	0,54	0,77	0,70	0,72	0,60	0,61	0,66	0,77	0,86	0,69	0,72	0,64	0,67	0,71	0,63	0,53	0,56	0,79	0,60
Q40	0,62	0,54	0,55	0,51	0,67	1,00	0,67	0,73	0,64	0,67	0,50	0,40	0,49	0,56	0,65	0,66	0,45	0,54	0,54	0,63	0,68	0,60	0,58	0,36	0,56	0,51
Q41	0,41	0,36	0,38	0,46	0,51	0,67	1,00	0,76	0,56	0,62	0,38	0,35	0,46	0,45	0,46	0,49	0,41	0,52	0,57	0,58	0,52	0,59	0,58	0,38	0,48	0,55
Q42	0,44	0,34	0,38	0,44	0,54	0,73	0,76	1,00	0,60	0,70	0,43	0,36	0,41	0,45	0,48	0,54	0,39	0,56	0,55	0,49	0,55	0,56	0,53	0,36	0,50	0,58
Q43	0,74	0,59	0,67	0,81	0,77	0,64	0,56	0,60	1,00	0,75	0,75	0,68	0,63	0,71	0,73	0,83	0,71	0,73	0,72	0,67	0,73	0,66	0,60	0,59	0,79	0,72
Q44	0,63	0,55	0,60	0,67	0,70	0,67	0,62	0,70	0,75	1,00	0,73	0,55	0,60	0,61	0,66	0,74	0,63	0,74	0,68	0,62	0,67	0,66	0,60	0,54	0,71	0,69
Q45	0,70	0,59	0,72	0,74	0,72	0,50	0,38	0,43	0,75	0,73	1,00	0,67	0,65	0,69	0,69	0,78	0,79	0,75	0,69	0,60	0,66	0,63	0,50	0,71	0,80	0,71
Q46	0,58	0,48	0,57	0,68	0,60	0,40	0,35	0,36	0,68	0,55	0,67	1,00	0,73	0,80	0,53	0,64	0,69	0,66	0,63	0,53	0,51	0,56	0,50	0,63	0,66	0,62
Q47	0,58	0,53	0,62	0,62	0,61	0,49	0,46	0,41	0,63	0,60	0,65	0,73	1,00	0,79	0,58	0,64	0,68	0,63	0,60	0,57	0,53	0,60	0,57	0,63	0,67	0,60
Q48	0,68	0,61	0,65	0,69	0,66	0,56	0,45	0,45	0,71	0,61	0,69	0,80	0,79	1,00	0,69	0,72	0,68	0,69	0,66	0,64	0,61	0,65	0,61	0,62	0,69	0,63
Q49	0,82	0,73	0,77	0,70	0,77	0,65	0,46	0,48	0,73	0,66	0,69	0,53	0,58	0,69	1,00	0,88	0,60	0,68	0,62	0,72	0,73	0,68	0,60	0,50	0,78	0,57
Q50	0,86	0,67	0,83	0,81	0,86	0,66	0,49	0,54	0,83	0,74	0,78	0,64	0,64	0,72	0,88	1,00	0,71	0,74	0,68	0,72	0,76	0,70	0,60	0,60	0,86	0,67
Q51	0,60	0,52	0,63	0,71	0,69	0,45	0,41	0,39	0,71	0,63	0,79	0,69	0,68	0,68	0,60	0,71	1,00	0,76	0,76	0,60	0,60	0,61	0,53	0,83	0,79	0,73
Q52	0,64	0,57	0,64	0,70	0,72	0,54	0,52	0,56	0,73	0,74	0,75	0,66	0,63	0,69	0,68	0,74	0,76	1,00	0,78	0,66	0,68	0,72	0,60	0,70	0,80	0,79
Q53	0,57	0,52	0,54	0,67	0,64	0,54	0,57	0,55	0,72	0,68	0,69	0,63	0,60	0,66	0,62	0,68	0,76	0,78	1,00	0,68	0,65	0,72	0,66	0,69	0,74	0,84
Q54	0,65	0,63	0,64	0,61	0,67	0,63	0,58	0,49	0,67	0,62	0,60	0,53	0,57	0,64	0,72	0,72	0,60	0,66	0,68	1,00	0,79	0,80	0,81	0,51	0,70	0,63
Q55	0,69	0,59	0,64	0,66	0,71	0,68	0,52	0,55	0,73	0,67	0,66	0,51	0,53	0,61	0,73	0,76	0,60	0,68	0,65	0,79	1,00	0,75	0,67	0,49	0,70	0,63
Q56	0,60	0,61	0,62	0,63	0,63	0,60	0,59	0,56	0,66	0,66	0,63	0,56	0,60	0,65	0,68	0,70	0,61	0,72	0,72	0,80	0,75	1,00	0,83	0,59	0,74	0,73
Q57	0,53	0,57	0,53	0,52	0,53	0,58	0,58	0,53	0,60	0,60	0,50	0,50	0,57	0,61	0,60	0,60	0,53	0,60	0,66	0,81	0,67	0,83	1,00	0,49	0,63	0,65
Q58	0,51	0,46	0,57	0,59	0,56	0,36	0,38	0,36	0,59	0,54	0,71	0,63	0,63	0,62	0,50	0,60	0,83	0,70	0,69	0,51	0,49	0,59	0,49	1,00	0,74	0,74
Q59	0,75	0,64	0,79	0,77	0,79	0,56	0,48	0,50	0,79	0,71	0,80	0,66	0,67	0,69	0,78	0,86	0,79	0,80	0,74	0,70	0,70	0,74	0,63	0,74	1,00	0,77
Q60	0,53	0,51	0,55	0,66	0,60	0,51	0,55	0,58	0,72	0,69	0,71	0,62	0,60	0,63	0,57	0,67	0,73	0,79	0,84	0,63	0,63	0,73	0,65	0,74	1,00	1,00

NO. CASES. 300

Anexo XII: Matrix de correlação (análise fatorial confirmatório) – Campus Pedra Branca

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q32	0,60	0,53	0,49	0,38	0,45	0,41	0,36	0,35	0,42	0,41	0,42	0,43	0,44	0,49	0,52	0,33	0,40	0,36	0,47	0,47	0,51	0,38	0,39	0,47	0,40
Q33	0,46	0,51	0,44	0,36	0,38	0,32	0,25	0,28	0,33	0,37	0,34	0,37	0,38	0,43	0,42	0,28	0,37	0,28	0,40	0,36	0,41	0,36	0,34	0,38	0,35
Q34	1,00	0,75	0,80	0,72	0,81	0,69	0,57	0,52	0,80	0,66	0,63	0,67	0,69	0,84	0,88	0,60	0,60	0,59	0,67	0,71	0,72	0,60	0,57	0,75	0,64
Q35	0,41	0,65	0,39	0,35	0,34	0,31	0,20	0,20	0,36	0,30	0,39	0,42	0,45	0,40	0,38	0,27	0,31	0,25	0,37	0,37	0,40	0,33	0,31	0,36	0,35
Q36	0,75	1,00	0,71	0,60	0,62	0,50	0,42	0,35	0,58	0,46	0,54	0,62	0,67	0,69	0,68	0,49	0,47	0,42	0,56	0,56	0,56	0,45	0,47	0,63	0,49
Q37	0,80	0,71	1,00	0,75	0,82	0,64	0,52	0,53	0,73	0,67	0,67	0,65	0,64	0,79	0,84	0,60	0,59	0,55	0,70	0,67	0,67	0,56	0,59	0,77	0,64
Q38	0,72	0,60	0,75	1,00	0,81	0,60	0,52	0,49	0,86	0,68	0,74	0,75	0,70	0,69	0,79	0,75	0,68	0,67	0,64	0,62	0,64	0,60	0,67	0,77	0,72
Q39	0,81	0,62	0,82	0,81	1,00	0,73	0,60	0,60	0,84	0,78	0,73	0,69	0,67	0,79	0,87	0,69	0,65	0,68	0,74	0,73	0,73	0,64	0,63	0,81	0,70
Q40	0,69	0,50	0,64	0,60	0,73	1,00	0,71	0,76	0,67	0,73	0,58	0,51	0,55	0,66	0,67	0,49	0,57	0,54	0,58	0,62	0,68	0,59	0,49	0,62	0,60
Q41	0,57	0,42	0,52	0,52	0,60	0,71	1,00	0,74	0,57	0,67	0,45	0,42	0,48	0,58	0,56	0,38	0,52	0,44	0,54	0,56	0,58	0,53	0,39	0,51	0,52
Q42	0,52	0,35	0,53	0,49	0,60	0,76	0,74	1,00	0,59	0,75	0,49	0,40	0,48	0,52	0,55	0,43	0,54	0,55	0,55	0,59	0,60	0,57	0,44	0,53	0,57
Q43	0,80	0,58	0,73	0,86	0,84	0,67	0,57	0,59	1,00	0,77	0,72	0,76	0,72	0,75	0,84	0,71	0,70	0,72	0,68	0,73	0,72	0,65	0,65	0,77	0,74
Q44	0,66	0,46	0,67	0,68	0,78	0,73	0,67	0,75	0,77	1,00	0,71	0,61	0,63	0,66	0,75	0,64	0,75	0,72	0,68	0,68	0,70	0,66	0,65	0,74	0,76
Q45	0,63	0,54	0,67	0,74	0,73	0,58	0,45	0,49	0,72	0,71	1,00	0,74	0,70	0,66	0,75	0,83	0,74	0,76	0,65	0,66	0,68	0,60	0,82	0,81	0,79
Q46	0,67	0,62	0,65	0,75	0,69	0,51	0,42	0,40	0,76	0,61	0,74	1,00	0,80	0,68	0,74	0,70	0,59	0,64	0,58	0,61	0,62	0,55	0,63	0,76	0,68
Q47	0,68	0,62	0,64	0,70	0,67	0,55	0,48	0,48	0,72	0,63	0,70	0,80	1,00	0,78	0,75	0,70	0,63	0,67	0,63	0,68	0,67	0,65	0,67	0,76	0,70
Q48	0,69	0,67	0,67	0,74	0,70	0,52	0,47	0,43	0,75	0,61	0,69	0,84	0,78	1,00	0,77	0,68	0,62	0,62	0,65	0,65	0,66	0,61	0,63	0,77	0,66
Q49	0,84	0,69	0,79	0,69	0,79	0,66	0,58	0,52	0,73	0,66	0,68	0,70	0,77	1,00	0,85	0,61	0,62	0,57	0,70	0,71	0,73	0,64	0,58	0,76	0,64
Q50	0,88	0,68	0,84	0,79	0,87	0,67	0,56	0,55	0,84	0,75	0,75	0,74	0,75	0,76	1,00	0,72	0,71	0,70	0,75	0,77	0,76	0,67	0,68	0,86	0,74
Q51	0,60	0,49	0,60	0,75	0,69	0,49	0,38	0,43	0,71	0,64	0,83	0,70	0,68	0,61	0,72	1,00	0,74	0,81	0,64	0,63	0,63	0,59	0,85	0,77	0,79
Q52	0,60	0,47	0,59	0,68	0,65	0,57	0,52	0,54	0,70	0,75	0,74	0,59	0,63	0,62	0,71	0,74	1,00	0,78	0,67	0,65	0,70	0,66	0,70	0,71	0,84
Q53	0,59	0,42	0,55	0,67	0,68	0,54	0,44	0,55	0,72	0,72	0,76	0,64	0,67	0,62	0,57	0,81	0,78	1,00	0,69	0,68	0,70	0,69	0,78	0,77	0,85
Q54	0,67	0,56	0,70	0,64	0,74	0,58	0,54	0,55	0,68	0,68	0,65	0,58	0,63	0,65	0,70	0,64	0,67	0,69	1,00	0,79	0,78	0,76	0,59	0,74	0,70
Q55	0,71	0,56	0,67	0,62	0,73	0,62	0,56	0,59	0,73	0,68	0,66	0,61	0,68	0,65	0,71	0,77	0,63	0,68	0,79	1,00	0,78	0,72	0,60	0,75	0,73
Q56	0,72	0,56	0,67	0,64	0,73	0,68	0,58	0,60	0,72	0,70	0,68	0,62	0,67	0,66	0,73	0,76	0,63	0,70	0,78	0,78	1,00	0,85	0,64	0,75	0,74
Q57	0,60	0,45	0,56	0,60	0,64	0,59	0,53	0,57	0,65	0,66	0,60	0,55	0,65	0,64	0,67	0,59	0,66	0,69	0,76	0,72	0,85	1,00	0,57	0,68	0,69
Q58	0,57	0,47	0,59	0,67	0,63	0,49	0,39	0,44	0,65	0,65	0,82	0,63	0,67	0,63	0,58	0,85	0,70	0,78	0,59	0,60	0,64	0,57	1,00	0,79	0,78
Q59	0,75	0,63	0,77	0,77	0,81	0,62	0,51	0,53	0,77	0,74	0,81	0,76	0,76	0,76	0,86	0,77	0,71	0,77	0,74	0,75	0,75	0,68	0,79	1,00	0,81
Q60	0,64	0,49	0,64	0,72	0,70	0,60	0,52	0,57	0,74	0,76	0,79	0,68	0,70	0,66	0,74	0,79	0,84	0,85	0,70	0,73	0,74	0,69	0,78	0,81	1,00

NO. CASES. 300

Anexo XIII: Matrix de correlação (análise fatorial confirmatório) – Geral Unisul

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,62	0,56	0,45	0,36	0,41	0,34	0,30	0,27	0,40	0,38	0,42	0,34	0,37	0,41	0,51	0,53	0,29	0,40	0,34	0,43	0,43	0,48	0,38	0,35	0,47	0,38
Q33	0,50	0,53	0,43	0,35	0,34	0,28	0,25	0,37	0,32	0,41	0,34	0,41	0,42	0,45	0,44	0,44	0,31	0,41	0,33	0,42	0,36	0,47	0,39	0,38	0,43	0,39
Q34	1,00	0,76	0,78	0,70	0,78	0,45	0,41	0,72	0,61	0,66	0,59	0,56	0,65	0,84	0,87	0,53	0,60	0,56	0,68	0,67	0,65	0,56	0,48	0,76	0,60	
Q35	0,40	0,60	0,34	0,30	0,25	0,30	0,22	0,32	0,27	0,31	0,33	0,37	0,38	0,37	0,36	0,21	0,28	0,24	0,38	0,29	0,36	0,33	0,25	0,31	0,31	
Q36	0,76	1,00	0,67	0,59	0,60	0,46	0,35	0,30	0,62	0,49	0,58	0,55	0,65	0,70	0,71	0,47	0,50	0,43	0,59	0,53	0,57	0,48	0,43	0,65	0,49	
Q37	0,78	0,67	1,00	0,74	0,82	0,56	0,42	0,40	0,70	0,62	0,72	0,61	0,59	0,67	0,79	0,83	0,56	0,59	0,70	0,63	0,62	0,54	0,52	0,76	0,60	
Q38	0,70	0,59	0,74	1,00	0,80	0,54	0,44	0,84	0,68	0,75	0,73	0,64	0,74	0,72	0,80	0,70	0,69	0,69	0,69	0,65	0,65	0,60	0,59	0,78	0,70	
Q39	0,78	0,60	0,82	0,80	1,00	0,66	0,51	0,54	0,82	0,75	0,76	0,60	0,69	0,79	0,86	0,66	0,67	0,67	0,73	0,70	0,69	0,60	0,56	0,81	0,67	
Q40	0,59	0,46	0,56	0,54	0,66	1,00	0,67	0,75	0,62	0,73	0,56	0,43	0,49	0,65	0,65	0,42	0,60	0,54	0,64	0,64	0,66	0,58	0,41	0,58	0,60	
Q41	0,45	0,35	0,42	0,44	0,51	0,67	1,00	0,72	0,51	0,65	0,44	0,38	0,45	0,41	0,48	0,50	0,35	0,52	0,47	0,52	0,58	0,54	0,31	0,47	0,51	
Q42	0,41	0,30	0,40	0,44	0,54	0,75	0,72	1,00	0,56	0,73	0,47	0,35	0,42	0,39	0,48	0,51	0,35	0,58	0,55	0,52	0,59	0,53	0,34	0,46	0,57	
Q43	0,72	0,62	0,70	0,84	0,82	0,62	0,51	0,56	1,00	0,75	0,78	0,76	0,67	0,76	0,74	0,84	0,71	0,73	0,74	0,72	0,73	0,65	0,61	0,82	0,77	
Q44	0,61	0,49	0,62	0,68	0,75	0,73	0,65	0,73	0,75	1,00	0,71	0,57	0,59	0,61	0,67	0,74	0,60	0,74	0,71	0,70	0,67	0,70	0,66	0,55	0,72	0,73
Q45	0,66	0,58	0,72	0,75	0,76	0,56	0,44	0,47	0,78	0,71	1,00	0,73	0,66	0,74	0,70	0,79	0,81	0,73	0,72	0,71	0,67	0,70	0,62	0,75	0,82	0,75
Q46	0,59	0,56	0,61	0,73	0,65	0,43	0,38	0,35	0,76	0,57	0,73	1,00	0,77	0,84	0,60	0,67	0,71	0,61	0,64	0,61	0,59	0,62	0,57	0,64	0,73	0,66
Q47	0,56	0,55	0,59	0,64	0,60	0,49	0,45	0,42	0,67	0,59	0,66	0,77	1,00	0,77	0,61	0,66	0,65	0,64	0,63	0,62	0,60	0,65	0,64	0,62	0,70	0,64
Q48	0,65	0,65	0,67	0,74	0,69	0,49	0,41	0,39	0,76	0,61	0,74	0,84	0,77	1,00	0,71	0,74	0,70	0,67	0,62	0,68	0,64	0,67	0,62	0,65	0,76	0,69
Q49	0,84	0,70	0,79	0,72	0,79	0,65	0,48	0,48	0,74	0,67	0,60	0,61	0,71	1,00	0,87	0,87	0,55	0,63	0,57	0,72	0,69	0,69	0,59	0,49	0,76	0,62
Q50	0,87	0,71	0,83	0,80	0,86	0,65	0,50	0,51	0,84	0,74	0,79	0,67	0,66	0,74	0,87	1,00	0,66	0,70	0,68	0,78	0,74	0,73	0,65	0,60	0,86	0,71
Q51	0,53	0,47	0,56	0,70	0,66	0,42	0,35	0,35	0,71	0,60	0,81	0,71	0,65	0,70	0,55	0,66	1,00	0,72	0,76	0,64	0,59	0,65	0,58	0,84	0,76	0,78
Q52	0,60	0,50	0,59	0,69	0,67	0,60	0,52	0,58	0,73	0,74	0,73	0,61	0,64	0,67	0,63	0,70	0,72	1,00	0,78	0,72	0,69	0,77	0,69	0,65	0,74	0,81
Q53	0,56	0,43	0,53	0,69	0,67	0,54	0,47	0,55	0,74	0,71	0,72	0,64	0,63	0,62	0,57	0,68	0,76	0,78	1,00	0,68	0,67	0,74	0,70	0,72	0,74	0,86
Q54	0,68	0,59	0,70	0,69	0,73	0,64	0,57	0,52	0,72	0,70	0,71	0,61	0,62	0,68	0,72	0,78	0,64	0,72	0,68	1,00	0,80	0,82	0,80	0,58	0,79	0,72
Q55	0,67	0,53	0,63	0,65	0,70	0,64	0,52	0,56	0,73	0,67	0,67	0,59	0,60	0,64	0,69	0,74	0,59	0,69	0,67	0,80	1,00	0,77	0,70	0,56	0,76	0,72
Q56	0,65	0,57	0,62	0,65	0,69	0,66	0,58	0,59	0,73	0,70	0,62	0,65	0,67	0,69	0,73	0,65	0,65	0,77	0,74	0,82	0,77	1,00	0,85	0,65	0,78	0,78
Q57	0,56	0,48	0,54	0,60	0,60	0,58	0,54	0,53	0,65	0,66	0,62	0,57	0,64	0,62	0,59	0,65	0,58	0,69	0,70	0,80	0,70	0,85	1,00	0,58	0,69	0,72
Q58	0,48	0,43	0,52	0,59	0,56	0,41	0,31	0,34	0,61	0,55	0,64	0,62	0,65	0,49	0,60	0,60	0,84	0,65	0,72	0,58	0,56	0,65	0,58	1,00	0,75	0,77
Q59	0,76	0,65	0,76	0,78	0,81	0,58	0,47	0,46	0,82	0,72	0,82	0,73	0,70	0,76	0,86	0,76	0,74	0,74	0,79	0,76	0,78	0,78	0,69	0,75	1,00	0,80
Q60	0,60	0,49	0,60	0,70	0,67	0,60	0,51	0,57	0,77	0,73	0,75	0,66	0,64	0,69	0,62	0,71	0,78	0,81	0,86	0,72	0,72	0,78	0,72	0,77	0,80	1,00
NO. CASES.																								330		

Anexo XIV: Matrix de correlação (análise fatorial confirmatório) – Campus Tubarão

Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60	
Q32	0,49	0,44	0,42	0,33	0,32	0,27	0,30	0,23	0,32	0,37	0,38	0,32	0,35	0,40	0,43	0,43	0,21	0,33	0,25	0,35	0,30	0,44	0,33	0,25	0,41	0,31
Q33	0,42	0,47	0,43	0,30	0,28	0,29	0,32	0,27	0,32	0,34	0,39	0,33	0,42	0,42	0,37	0,42	0,27	0,37	0,28	0,36	0,33	0,49	0,40	0,35	0,42	0,34
Q34	1,00	0,65	0,79	0,69	0,78	0,59	0,38	0,43	0,70	0,62	0,61	0,50	0,45	0,59	0,80	0,85	0,45	0,56	0,45	0,57	0,67	0,53	0,44	0,36	0,72	0,49
Q35	0,32	0,60	0,29	0,28	0,25	0,38	0,35	0,33	0,27	0,30	0,26	0,35	0,34	0,33	0,34	0,34	0,23	0,32	0,27	0,36	0,32	0,39	0,35	0,22	0,29	0,30
Q36	0,65	1,00	0,62	0,50	0,52	0,51	0,34	0,35	0,52	0,50	0,54	0,46	0,51	0,56	0,68	0,63	0,42	0,50	0,36	0,56	0,58	0,57	0,48	0,35	0,61	0,42
Q37	0,79	0,62	1,00	0,72	0,82	0,56	0,41	0,42	0,67	0,63	0,68	0,54	0,52	0,63	0,76	0,82	0,54	0,61	0,48	0,60	0,63	0,59	0,46	0,46	0,75	0,53
Q38	0,69	0,50	0,72	1,00	0,74	0,44	0,38	0,40	0,81	0,64	0,71	0,64	0,57	0,67	0,64	0,79	0,63	0,62	0,55	0,53	0,61	0,55	0,47	0,55	0,73	0,62
Q39	0,78	0,52	0,82	0,74	1,00	0,64	0,47	0,50	0,73	0,68	0,66	0,53	0,51	0,60	0,75	0,85	0,60	0,65	0,55	0,59	0,67	0,56	0,43	0,47	0,76	0,60
Q40	0,59	0,51	0,56	0,44	0,64	1,00	0,65	0,73	0,56	0,67	0,47	0,31	0,41	0,47	0,65	0,66	0,38	0,52	0,49	0,62	0,65	0,58	0,50	0,25	0,52	0,47
Q41	0,38	0,34	0,41	0,38	0,47	0,65	1,00	0,74	0,51	0,64	0,39	0,33	0,43	0,46	0,45	0,50	0,35	0,53	0,55	0,57	0,52	0,58	0,54	0,31	0,45	0,55
Q42	0,43	0,35	0,42	0,40	0,50	0,73	0,74	1,00	0,56	0,70	0,41	0,29	0,35	0,38	0,46	0,57	0,33	0,52	0,47	0,55	0,57	0,51	0,47	0,25	0,45	0,48
Q43	0,70	0,52	0,67	0,81	0,73	0,56	0,51	0,56	1,00	0,75	0,70	0,65	0,56	0,69	0,67	0,83	0,63	0,69	0,66	0,61	0,70	0,61	0,54	0,54	0,77	0,71
Q44	0,62	0,50	0,63	0,64	0,68	0,67	0,64	0,70	0,75	1,00	0,69	0,48	0,55	0,62	0,66	0,77	0,61	0,74	0,70	0,63	0,67	0,64	0,59	0,49	0,71	0,69
Q45	0,61	0,54	0,68	0,71	0,66	0,47	0,39	0,41	0,70	0,69	1,00	0,62	0,64	0,68	0,63	0,75	0,77	0,73	0,63	0,56	0,60	0,60	0,48	0,72	0,78	0,69
Q46	0,50	0,46	0,54	0,64	0,53	0,31	0,33	0,29	0,65	0,48	0,62	1,00	0,74	0,80	0,47	0,60	0,62	0,59	0,59	0,46	0,46	0,50	0,44	0,60	0,67	0,63
Q47	0,45	0,51	0,52	0,57	0,51	0,41	0,43	0,35	0,56	0,55	0,64	0,74	1,00	0,78	0,52	0,57	0,61	0,60	0,56	0,48	0,46	0,59	0,49	0,59	0,63	0,60
Q48	0,59	0,56	0,63	0,67	0,60	0,47	0,46	0,38	0,69	0,62	0,68	0,80	0,78	1,00	0,65	0,70	0,61	0,63	0,58	0,60	0,56	0,61	0,57	0,58	0,70	0,64
Q49	0,80	0,68	0,76	0,64	0,75	0,65	0,45	0,46	0,67	0,66	0,63	0,47	0,52	0,65	1,00	0,85	0,50	0,62	0,52	0,67	0,73	0,63	0,51	0,37	0,72	0,53
Q50	0,85	0,63	0,82	0,79	0,85	0,66	0,50	0,57	0,83	0,77	0,75	0,60	0,57	0,70	0,85	1,00	0,65	0,70	0,61	0,68	0,74	0,67	0,55	0,52	0,83	0,67
Q51	0,45	0,42	0,54	0,63	0,60	0,38	0,35	0,33	0,63	0,61	0,77	0,62	0,61	0,61	0,50	0,65	1,00	0,74	0,73	0,53	0,53	0,57	0,49	0,83	0,75	0,75
Q52	0,56	0,50	0,61	0,62	0,65	0,52	0,53	0,52	0,69	0,74	0,73	0,59	0,60	0,63	0,62	0,70	0,74	1,00	0,77	0,63	0,62	0,70	0,56	0,69	0,77	0,79
Q53	0,45	0,36	0,48	0,55	0,55	0,49	0,55	0,47	0,66	0,70	0,63	0,59	0,56	0,58	0,52	0,61	0,73	0,77	1,00	0,63	0,60	0,69	0,64	0,65	0,68	0,85
Q54	0,57	0,56	0,60	0,53	0,59	0,62	0,57	0,55	0,61	0,63	0,56	0,46	0,48	0,60	0,67	0,68	0,53	0,63	0,63	1,00	0,76	0,78	0,78	0,45	0,64	0,63
Q55	0,67	0,58	0,63	0,61	0,67	0,65	0,52	0,57	0,70	0,67	0,60	0,46	0,46	0,56	0,73	0,74	0,53	0,62	0,60	0,76	1,00	0,70	0,65	0,43	0,69	0,64
Q56	0,53	0,57	0,59	0,55	0,56	0,58	0,58	0,51	0,61	0,64	0,60	0,50	0,59	0,61	0,63	0,67	0,57	0,70	0,69	0,78	0,70	1,00	0,82	0,56	0,69	0,74
Q57	0,44	0,48	0,46	0,47	0,43	0,50	0,54	0,47	0,54	0,59	0,48	0,44	0,49	0,57	0,51	0,55	0,49	0,56	0,64	0,78	0,65	0,82	1,00	0,46	0,55	0,65
Q58	0,36	0,35	0,46	0,55	0,47	0,25	0,31	0,25	0,54	0,49	0,72	0,60	0,59	0,58	0,37	0,52	0,83	0,69	0,65	0,45	0,43	0,56	0,46	1,00	0,72	0,75
Q59	0,72	0,61	0,75	0,73	0,76	0,52	0,45	0,45	0,77	0,71	0,78	0,67	0,63	0,70	0,72	0,83	0,75	0,77	0,68	0,64	0,69	0,69	0,55	0,72	1,00	0,78
Q60	0,49	0,42	0,53	0,62	0,60	0,47	0,55	0,48	0,71	0,69	0,69	0,63	0,60	0,64	0,53	0,67	0,75	0,79	0,85	0,63	0,64	0,74	0,65	0,75	0,78	1,00

NO. CASES. 330

Anexo XV: Matrix de correlação (análise fatorial confirmatório) – Campus Pedra Branca

	Q34	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56	Q57	Q58	Q59	Q60
Q32	0,50	0,44	0,47	0,38	0,42	0,36	0,22	0,27	0,46	0,37	0,38	0,39	0,35	0,37	0,42	0,47	0,30	0,46	0,32	0,36	0,42	0,50	0,30	0,32	0,49	0,39
Q33	0,34	0,41	0,38	0,38	0,33	0,29	0,13	0,11	0,36	0,26	0,42	0,42	0,34	0,43	0,38	0,34	0,35	0,40	0,37	0,37	0,33	0,45	0,30	0,42	0,40	0,38
Q34	1,00	0,70	0,76	0,67	0,78	0,60	0,31	0,35	0,74	0,56	0,62	0,52	0,54	0,64	0,79	0,87	0,47	0,55	0,40	0,58	0,68	0,58	0,45	0,40	0,69	0,42
Q35	0,38	0,60	0,34	0,36	0,35	0,47	0,30	0,31	0,34	0,34	0,41	0,35	0,38	0,38	0,40	0,38	0,26	0,36	0,26	0,45	0,41	0,39	0,31	0,29	0,37	0,33
Q36	0,70	1,00	0,65	0,53	0,62	0,54	0,29	0,25	0,53	0,39	0,52	0,45	0,55	0,59	0,70	0,64	0,35	0,46	0,34	0,54	0,57	0,52	0,46	0,35	0,59	0,37
Q37	0,76	0,65	1,00	0,69	0,79	0,56	0,34	0,35	0,66	0,57	0,61	0,54	0,53	0,62	0,75	0,82	0,51	0,55	0,42	0,59	0,57	0,57	0,45	0,40	0,70	0,43
Q38	0,67	0,53	0,69	1,00	0,76	0,53	0,34	0,38	0,86	0,64	0,75	0,74	0,62	0,73	0,67	0,78	0,68	0,65	0,59	0,56	0,63	0,58	0,47	0,61	0,77	0,61
Q39	0,78	0,62	0,79	0,76	1,00	0,65	0,41	0,43	0,77	0,65	0,70	0,61	0,62	0,70	0,78	0,85	0,60	0,62	0,52	0,63	0,70	0,60	0,48	0,50	0,78	0,51
Q40	0,60	0,54	0,56	0,53	0,65	1,00	0,60	0,66	0,58	0,65	0,58	0,44	0,53	0,53	0,64	0,65	0,42	0,56	0,47	0,62	0,68	0,59	0,44	0,39	0,59	0,49
Q41	0,31	0,29	0,34	0,34	0,41	0,60	1,00	0,69	0,42	0,59	0,38	0,25	0,44	0,40	0,36	0,41	0,41	0,52	0,47	0,54	0,48	0,53	0,51	0,39	0,44	0,49
Q42	0,35	0,25	0,35	0,38	0,43	0,66	0,69	1,00	0,43	0,70	0,41	0,27	0,35	0,32	0,38	0,44	0,34	0,56	0,47	0,44	0,44	0,46	0,40	0,32	0,42	0,45
Q43	0,74	0,53	0,66	0,86	0,77	0,58	0,42	0,43	1,00	0,69	0,76	0,75	0,66	0,75	0,69	0,82	0,71	0,71	0,66	0,64	0,71	0,66	0,56	0,65	0,81	0,69
Q44	0,56	0,39	0,57	0,64	0,65	0,65	0,59	0,70	0,69	1,00	0,70	0,49	0,52	0,56	0,59	0,70	0,64	0,77	0,66	0,62	0,58	0,63	0,55	0,58	0,71	0,68
Q45	0,62	0,52	0,61	0,75	0,70	0,58	0,38	0,41	0,76	0,70	1,00	0,69	0,64	0,69	0,65	0,75	0,80	0,77	0,64	0,60	0,64	0,68	0,52	0,76	0,83	0,69
Q46	0,52	0,45	0,54	0,74	0,61	0,44	0,25	0,27	0,75	0,49	0,69	1,00	0,70	0,80	0,58	0,64	0,65	0,58	0,57	0,48	0,50	0,53	0,44	0,62	0,68	0,59
Q47	0,54	0,55	0,53	0,62	0,62	0,53	0,44	0,35	0,66	0,52	0,64	0,70	1,00	0,75	0,64	0,66	0,61	0,59	0,55	0,57	0,60	0,61	0,51	0,55	0,68	0,56
Q48	0,64	0,59	0,62	0,73	0,70	0,53	0,40	0,32	0,75	0,56	0,69	0,80	0,75	1,00	0,72	0,73	0,62	0,63	0,55	0,62	0,62	0,62	0,53	0,58	0,72	0,58
Q49	0,79	0,70	0,75	0,67	0,78	0,64	0,36	0,38	0,69	0,59	0,65	0,58	0,64	0,72	1,00	0,85	0,50	0,56	0,49	0,67	0,68	0,65	0,54	0,43	0,73	0,50
Q50	0,87	0,64	0,82	0,78	0,85	0,65	0,41	0,44	0,82	0,70	0,75	0,64	0,66	0,73	0,85	1,00	0,64	0,66	0,55	0,69	0,71	0,68	0,55	0,53	0,83	0,58
Q51	0,47	0,35	0,51	0,68	0,60	0,42	0,41	0,34	0,71	0,64	0,80	0,65	0,61	0,62	0,50	0,64	1,00	0,75	0,75	0,55	0,55	0,62	0,53	0,88	0,79	0,77
Q52	0,55	0,46	0,55	0,65	0,62	0,56	0,52	0,56	0,71	0,77	0,77	0,58	0,59	0,63	0,56	0,66	0,75	1,00	0,77	0,64	0,64	0,71	0,58	0,73	0,75	0,79
Q53	0,40	0,34	0,42	0,59	0,52	0,47	0,47	0,47	0,66	0,66	0,64	0,57	0,55	0,55	0,49	0,55	0,75	0,77	1,00	0,65	0,60	0,72	0,67	0,72	0,69	0,84
Q54	0,58	0,54	0,59	0,56	0,63	0,62	0,54	0,44	0,64	0,62	0,60	0,48	0,57	0,62	0,67	0,69	0,55	0,64	0,65	1,00	0,77	0,76	0,78	0,48	0,66	0,66
Q55	0,68	0,57	0,57	0,63	0,70	0,68	0,48	0,44	0,71	0,58	0,64	0,50	0,60	0,62	0,68	0,71	0,55	0,64	0,60	0,77	1,00	0,75	0,65	0,48	0,70	0,59
Q56	0,58	0,52	0,57	0,58	0,60	0,59	0,53	0,46	0,66	0,63	0,68	0,53	0,61	0,62	0,65	0,68	0,62	0,71	0,72	0,76	0,75	1,00	0,79	0,59	0,74	0,74
Q57	0,45	0,46	0,45	0,47	0,48	0,44	0,51	0,40	0,56	0,55	0,52	0,44	0,51	0,53	0,54	0,55	0,53	0,58	0,67	0,78	0,65	0,79	1,00	0,49	0,63	0,64
Q58	0,40	0,35	0,40	0,61	0,50	0,39	0,39	0,32	0,65	0,58	0,76	0,62	0,55	0,58	0,43	0,53	0,88	0,73	0,72	0,48	0,48	0,59	0,49	1,00	0,72	0,76
Q59	0,69	0,59	0,70	0,77	0,78	0,59	0,44	0,42	0,81	0,71	0,83	0,68	0,68	0,72	0,73	0,83	0,79	0,75	0,69	0,66	0,70	0,74	0,63	0,72	1,00	0,73
Q60	0,42	0,37	0,43	0,61	0,51	0,49	0,49	0,45	0,69	0,68	0,69	0,59	0,56	0,58	0,50	0,58	0,77	0,79	0,84	0,66	0,59	0,74	0,64	0,76	0,73	1,00

NO. CASES. 330