

HUGO MUNETTI



EVIDÊNCIAS DE VALIDADE DE UM INSTRUMENTO DE
AVALIAÇÃO DOCENTE

ITATIBA

2007

HUGO MUSETTI

EVIDÊNCIAS DE VALIDADE DE UM INSTRUMENTO DE
AVALIAÇÃO DOCENTE

Dissertação apresentada ao Programa de Pós-
Graduação *Stricto Sensu* em Psicologia da
Universidade São Francisco para obtenção do
título de Mestre.

ORIENTADORA: PROF^a DR^a CLAUDETTE MARIA MEDEIROS VENDRAMINI

ITATIBA

2007

UNIVERSIDADE SÃO FRANCISCO

PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM PSICOLOGIA

MESTRADO

EVIDÊNCIAS DE VALIDADE DE UM INSTRUMENTO DE
AVALIAÇÃO DOCENTE

AUTOR: HUGO MUSETTI

ORIENTADORA: PROF^a DR^a CLAUDETTE MARIA MEDEIROS VENDRAMINI

Este exemplar corresponde à redação final da dissertação de
mestrado defendida por Hugo Musetti e aprovada pela
comissão examinadora. Data: 26 /06/ 2007.

COMISSÃO EXAMINADORA

Prof^a Dr^a Claudette Maria MedeirosVendramini (Orientadora)

Prof^a Dr^a Márcia Regina Ferreira de Brito

Prof^a Dr^a Alessandra Gotuzo Seabra Capovilla

Itatiba

2007

Dedicatória

À minha família,
especialmente pela paciência, apoio e incentivo, e porque muitas vezes estive ausente, para
que pudesse realizar as pesquisas. De modo especial aos meus pais Humberto e Maria de
Lourdes, pelo exemplo de amor incondicional, tenho orgulho de ser filho de vocês!

Agradecimentos

À minha orientadora, Profª Drª Claudette Maria Medeiros Vendramini, que confiou e acreditou em mim desde o primeiro dia, muito me ensinou, apoiou, e incentivou quando foi necessário, um exemplo de educadora, profissional e ser humano.

À coordenação, aos colaboradores e professores do Programa de Pós-Graduação em Psicologia da Universidade São Francisco, especialmente a Profª Drª Maria Cristina Rodrigues Azevedo Joly e Profª Drª Alessandra Gotuzzo Seabra Capovilla, pelas valiosas contribuições que fizeram ao trabalho.

Aos meus colegas de turma, que muito me incentivaram e contribuíram para a construção do meu conhecimento e crescimento profissional.

À equipe da Comissão Própria de Avaliação Institucional, Profª Ms. Vânia Franciscon Vieira e Solange Aparecida da Silva Consolini pelo encorajamento.

À Profª Jaqueline Cristiane Stavis, pelo apoio e carinho que demonstra a cada dia, também pela força, paciência e por estar ao meu lado ao longo dessa caminhada.

Resumo

Musetti, Hugo (2007). *Evidências de validade de um instrumento de avaliação docente*. Dissertação de Mestrado, Programa de Pós-Graduação *Stricto Sensu* em Psicologia, Universidade São Francisco, Itatiba, SP.

O desempenho docente é um importante indicador de qualidade de ensino nas instituições educacionais, e sua avaliação faz parte de um processo permanente inserido em uma política de qualidade mais ampla, realizada externamente por órgãos governamentais e internamente por comissões próprias de avaliação. Nesse processo, é de fundamental importância utilizar instrumentos válidos e fidedignos para aferir o comportamento das pessoas que participam das instituições educacionais. Assim, pretende-se contribuir com o avanço em avaliação educacional investigando sobre a qualidade de medidas de desempenho docente em instituições de ensino superior com base em técnicas psicométricas de avaliação de instrumentos. Neste contexto, o objetivo do presente estudo foi verificar as propriedades psicométricas, validade e fidedignidade de um instrumento de avaliação docente realizada por discentes de uma universidade particular do interior do estado de São Paulo. Este instrumento é composto de quatro indicadores de desempenho docente: didática, conteúdo, processo de avaliação e relacionamento. O estudo foi realizado a partir de um banco de dados que contém informações de 628 docentes avaliados em 2005 por 5.568 discentes e de 618 docentes avaliados em 2006 por 6.072 discentes, distribuídos em 32 cursos de diferentes áreas de conhecimento de quatro câmpus universitários. Cada estudante avaliou espontaneamente os docentes das disciplinas que cursavam em 2005 e 2006. Pode-se concluir que o instrumento é unidimensional, e que 75,3% da variância total é explicada pelo primeiro componente da estrutura fatorial, valor considerado alto. Além disso, o instrumento é consistente internamente, com um coeficiente alfa de Cronbach igual a 0,890.

Palavras-chave: Educação Superior, avaliação estudantil, avaliação educacional, psicometria, desempenho docente.

Abstract

Musetti, Hugo (2007). *Validity evidences of a teaching evaluation instrument*. Master's degree dissertation, Graduate Studies in Psychology. San Francisco University, Itatiba, SP.

The teaching performance is a important indicator of the teaching quality in the educational institutions and its avaliation is part of a permanent process that belongs to a wider quality policy made externally by governamental sectors and internally by its avaliation comissions. In this process, the use of reliable instruments and with validity to measure the people's behavior that participate of the educational institutions is fundamental. So, the aim of this work is to contribute with the development of the educational avaliation investigating about the quality of the measures of the teaching performance in the universities based on psychometric technics of instrument avaliation. In this context, the objective of this study is to verify the psychometric properties, validity and reliability of an instrument of teaching avaliation in a private university in the State of Sao Paulo. This instrument is composed by four indicators of teaching performance: didactics, content, avaliation process and relationship. This study is based on information of about 628 teachers assessed by 5,568 students in 2005 and of 618 teachers assessed by 6,072 students in 2006 in 32 courses of different fields inside four university campi. Each student avaliated spontaneously the teachers of the disciplines they were attending in 2005 and 2006. It can be concluded that the instrument is unidimensional, and that 75.3% of the total variance are explained by the first component of the factorial structure, considered value high. Moreover, the instrument is consistent internally, with a Cronbach alpha coefficient equal the 0.890.

Keywords: college education, student's avaliation, educational avaliation, psychometry, teacher's performance.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS.....	xi
LISTA DE ANEXOS.....	xii
LISTA DE ABREVIATURAS.....	xiii
INTRODUÇÃO	1
SISTEMAS DE AVALIAÇÃO EDUCACIONAL E DE DESEMPENHO DOCENTE.....	10
A AVALIAÇÃO DE DESEMPENHO DOCENTE	19
PESQUISAS REFERENTES À AVALIAÇÃO DE DESEMPENHO DOCENTE.....	29
OBJETIVO GERAL	38
OBJETIVOS ESPECÍFICOS.....	38
MÉTODO.....	39
FORTE DE DADOS	39
MATERIAL	42
PROCEDIMENTO.....	46
ANÁLISE DE DADOS.....	49
RESULTADOS E DISCUSSÕES	52
CONCLUSÃO	62
BIBLIOGRAFIA.....	64
ÍNDICE REMISSIVO POR AUTOR.....	69
ANEXOS.....	71

LISTA DE FIGURAS

Figura 1. Distribuição do número de IES do Brasil no período de 1991 a 2005 por categoria administrativa	11
Figura 2. Distribuição do número de alunos inscritos nas IES do Brasil no período de 1991 a 2005	12
Figura 3. Distribuição do número de docentes em exercício no período de 1991 a 2005 por titulação	13
Figura 4. Paradigmas que inspiram as estratégias de avaliação da eficácia do desempenho docente segundo Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999).....	21
Figura 5. Esquema prognóstico-processo-produto adotado por Braskamp, Brandenburg & Cry (1994, citado por Tejedor & Monteiro, 1990)	22
Figura 6. Esquema proposto por Wilman (1988, citado por Tejedor & Monteiro, 1990) sobre as condutas de um professor	23
Figura 7. Referenciais técnicos e assistenciais identificados por Rodríguez Espinar (1988, citado por Tejedor & Monteiro, 1990) necessários na relação direta com a qualidade docente.....	23
Figura 8. Representação gráfica dos traços relacionados com aspectos da atividade docente segundo Tejedor e colaboradores (1988).....	24
Figura 10. Fluxo do processo de avaliação docente	47
Figura 11. Gráfico autovalor	57

Figura 12. Tela de autenticação do aluno	71
Figura 13. Tela de avaliação do aluno referente à Didática	71
Figura 14. Tela de avaliação do aluno referente ao Conteúdo	72
Figura 15. Tela de avaliação do aluno referente à Avaliação.....	72
Figura 16. Tela de avaliação do aluno referente ao Relacionamento.....	73

LISTA DE TABELAS

Tabela 1. Distribuição dos fatores de qualidade docente encontradas por diferentes autores segundo as dimensões de qualidade docente da presente pesquisa	32
Tabela 2. Distribuição do número de docentes avaliados por ano, semestre e gênero.....	39
Tabela 3. Distribuição do número de docentes avaliados por titulação, ano e semestre.....	40
Tabela 4. Distribuição do número de estudantes que responderam a avaliação docente por ano, semestre e gênero dos estudantes	40
Tabela 5. Distribuição do número de estudantes que responderam a avaliação docente por curso, ano e semestre	41
Tabela 6. Estatísticas descritivas dos itens da escala de avaliação docente de 15 itens - 2004	44
Tabela 7. Estatísticas descritivas da avaliação docente realizada pelo discente por ano, semestre, item do instrumento e câmpus	52
Tabela 8. Matriz de correlação da avaliação docente pelo discente realizada nos semestres de 2005 e 2006.....	54
Tabela 9. Comunalidade extraída da avaliação docente pelo discente realizada nos semestres e 2005 e de 2006	55
Tabela 10. Autovalor e variância explicada dos componentes da avaliação docente	55
Tabela 11. Distribuição das médias alcançadas pelos docentes atribuídas pelo aluno, estratificada de acordo com os percentis P_{95} e P_5	58

Tabela 13. Distribuição das médias alcançadas pelos docentes de acordo com o gênero do docente e do aluno	59
---	----

Tabela 13. Coeficiente de fidedignidade alfa de Cronbach da escala por ano e semestre ...	60
---	----

LISTA DE ANEXOS

Telas do formulário de avaliação docente	71
--	----

LISTA DE ABREVIATURAS

ACS	Assessoria de Comunicação Social do Ministério da Educação
BSRI	<i>Bem Sex-Role Inventory</i>
CONAES	Comissão Nacional de Avaliação da Educação Superior
CPA	Comissão Própria de Avaliação
CTAA	Comissão Técnica de Acompanhamento da Avaliação
IDES	Índice do Desenvolvimento do Ensino Superior
IES	Instituição de Ensino Superior
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
MEC	Ministério da Educação
TFQ	Questionário para Avaliação de Ensino

INTRODUÇÃO

A preocupação com a qualidade do ensino tem ocupado, nos últimos anos, cada vez mais um lugar no discurso dos dirigentes governamentais, dos mantenedores das instituições de ensino e da sociedade em geral. A avaliação do desempenho docente se assume como processo de investigação permanente inserida numa política de qualidade mais ampla para as Instituições de Ensino Superior (IES), conduzida externamente por órgãos governamentais e internamente por comissões técnicas. Quanto mais a sociedade evolui, mais variadas e urgentes se tornam as necessidades de se poder avaliar o comportamento das pessoas de uma forma mais precisa, pois, com base em tais avaliações, são tomadas decisões que afetam profundamente a vida dessas mesmas pessoas. Desta forma, faz-se necessário realizar avaliações que resultem em resultados confiáveis que auxiliem a tomada de decisão em processos educacionais (Pasquali, 2001).

A avaliação do docente tem o propósito de estimular reflexões ao corpo docente considerando os diferentes aspectos apontados pelos estudantes quanto à condução didático-pedagógica e o relacionamento mantido do professor com a turma ao longo do período letivo. Em geral, os resultados de uma avaliação docente possibilitam identificar possíveis problemas pedagógicos e comportamentais em situações vivenciadas pelos estudantes que poderiam comprometer o adequado desenvolvimento do processo de ensino-aprendizagem. Por outro lado, essas avaliações possibilitam identificar, pontos positivos do desempenho docente e que contribuem para um adequado ambiente de ensino. Em resumo, a avaliação dos professores pode servir como norteador de ações que favoreçam a construção da responsabilidade coletiva com o sistema educativo e o desenvolvimento dos profissionais ligados à educação (Duke & Stiggins, 1997 citado por

Fuentes-Medina & Sánchez, 1999), mas, para tanto, deve-se cuidar que essa avaliação seja válida e precisa, ressalta Fuentes-Medina e Sánchez (1999).

Ao realizar um diagnóstico das IES, seja este conduzido internamente pelas Comissões Próprias de Avaliação (CPA), seja externamente, pelo Ministério da Educação (MEC), por intermédio de seus institutos, secretarias e coordenações, o desempenho do corpo docente é considerado um importante indicador para aferir a qualidade do ensino ofertado, mesmo diante das limitações teóricas e práticas desse processo apontadas por Tejedor e Monteiro (1990). Conforme esses autores não existem um acordo unânime sobre as finalidades do ensino universitário, nem se dispõe de um modelo ideal de professor universalmente aceito que possa servir de marco de referência para o estudo avaliativo, o que se constitui em uma limitação teórica. Quanto às limitações de natureza prática, estas decorrem da dificuldade de eleger uma estratégia avaliativa adequada e válida que explicita com suficiente clareza a relação que se estabelece entre as distintas variáveis contempladas para um professor ideal.

Miller (1998, citado por Almeida, 2002) aponta que o corpo docente de uma universidade é a principal fonte do conhecimento acadêmico. Nesse sentido, muitas pesquisas têm sido conduzidas, no exterior em maior quantidade que no Brasil, a respeito da avaliação do desempenho dos docentes pelos estudantes. Dessa forma, o presente estudo surgiu basicamente da necessidade de realizar pesquisas envolvendo instrumentos de avaliação docente cujos resultados possam contribuir para a melhoria da qualidade do corpo docente nas IES do Brasil e, conseqüentemente, colaborar para o aprimoramento da qualidade do ensino, mesmo que de forma modesta. Além disso, os métodos e a qualidade dos instrumentos utilizados no processo de avaliação das IES ainda são muito questionados

pela comunidade acadêmica e têm sido foco de atenção de pesquisadores educacionais brasileiros.

Nesse sentido, são utilizados vários métodos que possibilitam verificar tais qualidades, como as descritas por Tejedor e Monteiro (1990) e Fuentes-Medina e Sánchez (1999): observações sistemáticas por meio das gravações em vídeo das aulas ministradas pelo docente, auto-avaliações docentes, provas padronizadas, avaliação de juízes, avaliação pelos pares e as avaliações realizadas pelos estudantes. Todas essas técnicas podem contribuir para melhorar a qualidade do trabalho docente. Quando se utiliza a técnica de observação sistemática, se obtêm registros válidos e fidedignos do desempenho docente em sala de aula. Esta é uma prática que pode contribuir para a melhoria do trabalho docente, mas é importante que as gravações realizadas sejam discutidas pelo professor, por um especialista e por um colega. No entanto, para que bons resultados sejam alcançados a presença da câmera em sala de aula não pode interferir no desempenho do docente.

Como destacam esses autores, a técnica de auto-avaliação é aquela em que o sujeito avaliado e o avaliador coincidem. Essa técnica consiste em analisar as reflexões feitas pelos próprios docentes sobre seu trabalho, e podem ser norteados pelos objetivos, materiais, procedimentos e atividades conduzidas em sala. Outra técnica de avaliação docente que pode ser considerada como uma forma indireta de avaliar o desempenho docente é a utilização das provas padronizadas aplicadas aos estudantes. O bom desempenho discente pode estar positivamente associado ao bom desempenho docente, embora este desempenho não dependa exclusivamente das atividades que o professor realiza em sala de aula, pois existem outras variáveis que poderiam interferir neste processo.

Como citam Fuentes-Medina e Sánchez (1999), as avaliações realizadas pelos pares, acompanhadas ou não da observação sistemática do comportamento docente em sala de

aula, são úteis para investigar aspectos do trabalho desenvolvido em classe, o profissionalismo do professor e o domínio do conteúdo. Em geral os pares podem estimar com bastante exatidão a competência de um professor, não tanto no que se refere a sua conduta em sala, mas sim, quanto à bagagem de conteúdo e sua preparação pedagógica e científica.

As avaliações realizadas pelos estudantes, na opinião de Tejedor e Monteiro (1990), é o melhor critério para aferir a qualidade do ensino, em virtude dos estudantes serem fiéis observadores. Sem dúvida, os critérios mais empregados nas IES para aferir a qualidade do docente em sala são os que compõem as opiniões dos estudantes (Moomaw, 1977 citado por Tejedor, Seijas & Rodríguez, 1988). Essa abordagem pode compreender entrevistas em grupo, comentários escritos ou questionários padronizados.

Informações interessantes sobre o trabalho docente podem ser obtidas a partir de entrevistas em grupo com os estudantes, mas sua condução deve ser realizada por profissionais com formação adequada em pesquisas e análise de conteúdo. Quando se analisam os comentários feitos pelos estudantes sobre o professor nas entrevistas é possível obter um valioso diagnóstico, mas a categorização destas informações coletadas é mais difícil e geralmente demanda um longo tempo. Outra forma empregada são os questionários estruturados e estes são, normalmente, os mais utilizados pelas IES para a coleta de informações. Em geral, são instrumentos estruturados em papel ou de forma eletrônica, compostos por um determinado número de perguntas, normalmente com o emprego de escalas do tipo *likert*, as quais fornecem direções sobre as respostas dos estudantes a cada afirmação, positiva ou negativa, em relação a determinados atributos relevantes do ponto de vista da eficácia docente.

Muitas vezes, com o objetivo de ofertar ao estudante um espaço de expressão para seus anseios acadêmicos, além dos instrumentos avaliativos aplicados nas IES são coletados outros dados de caráter qualitativo, que, na maioria das vezes, complementam as informações coletadas com a aplicação de escalas do tipo *likert* (Lowman, 2004).

Decorrente do grande número de investigações da avaliação discente sobre o desempenho docente, vários estudos estão voltados aos possíveis vieses que essa avaliação poderia ter, como por exemplo, a influência de gênero na avaliação do estudante, atribuições de notas altas pelo professor, ensino facilitado, interação no relacionamento professor–aluno, tamanho da turma, responsabilidade dos estudantes com a aprendizagem, até estudos sobre a influência da personalidade do professor na avaliação e o momento da aplicação do instrumento (Armstrong, 1983; Marsh, 1984; Greenwald & Gillmore, 1997; entre outros).

Pesquisas citadas por Tejedor e Monteiro (1990) e Centra e Gaubatz (1999) mostram que não existem diferenças significativas entre as avaliações realizadas por alunos e alunas. Outras denotam que as alunas são mais críticas que os alunos, tendendo a valorar melhor as professoras que os professores. Outras pesquisas indicam que as alunas valoram melhor os professores, em determinadas dimensões, que os alunos. Tais resultados são difíceis de interpretar e apresentam resultados muitas vezes contraditórios, pois a avaliação pode sofrer influência do papel social que, em ambos os sexos, é influenciado pela cultura.

Tejedor e Monteiro (1990) e Lowman (2004) concordam que existe certo sentimento entre o corpo docente de que alunos com boas notas atribuem pontuações mais altas para os seus docentes do que alunos com notas mais baixas, ou que professores que não apresentam bom desempenho tentem “comprar” alunos atribuindo-lhes notas mais altas. Esses fatores poderiam comprometer a validade do instrumento, no entanto, segundo

esses autores, a literatura indica que não existe uma conclusão sobre as relações entre esses fatores e que a falta de acordo entre os resultados se deve a razões de ordem metodológica. Em alguns estudos os procedimentos de amostragem aplicados não permitem a generalização dos resultados, em outros, as unidades de análise são diferentes nos diversos estudos (Lowman, 2004; Tejedor & Monteiro, 1990).

Lowman (2004), verificou que alguns professores acreditam que as boas avaliações atribuídas pelos estudantes devem ser reflexo de um ensino facilitado, e que se cultiva a esperança de que os estudantes que os avaliam como ruins irão, um dia, valorizá-los mais do que aqueles professores que eles acharam satisfatórios no momento da avaliação. O rigor mantido na condução das aulas e a busca de um bom desenvolvimento e aprofundamento dos conteúdos das disciplinas ministradas poderiam não estar associados à competência técnica do professor. Alguns pesquisadores (Abbot & Perkins, 1978; Reardon & Waters, 1979 citado por Lowman, 2004) expressam-se contrários a essa posição, pautados em resultados de pesquisa que evidenciam que os estudantes consideram a qualidade do relacionamento existente entre o professor e o aluno em segundo lugar de importância, depois da habilidade do professor expor de forma clara e envolvente, os conteúdos das disciplinas ministradas. Cunha (1999) contrapõe esta idéia de que um bom relacionamento entre professores e estudantes, os chamados “professores bonzinhos”, refletem em altos conceitos atribuídos pelos estudantes a esses docentes constituindo-se em avaliações tendenciosas. Segundo esse autor, os estudantes valorizam os professores exigentes, que cobram dos alunos a participação e a realização de tarefas.

Um fator que poderia interferir na qualidade do desempenho docente refere-se ao relacionamento existente entre professor–aluno que pode ser intensificado quando o trabalho do professor é conduzido em pequenos grupos, no entanto, Greenwald e Gillmore

(1997) e Tejedor e colaboradores (1988) relatam que ainda não se chegou a uma conclusão entre o trabalho do docente e o número de alunos em sala de aula

Lowman (2004) ressalta que para alguns docentes, os conceitos atribuídos pelos estudantes são inconsistentes em virtude da falta de maturidade dos mesmos em perceber que determinadas ações pedagógicas muitas vezes se fazem necessárias para o próprio aprendizado do alunado. Esses docentes enfatizam, ainda, que seus estudantes somente têm condições de avaliá-los após atuarem no mercado de trabalho desempenhando atividades em áreas correlacionadas com a disciplina cursada. Segundo esse autor, estudos retratam que as avaliações feitas do docente, vários anos depois da graduação, demonstram ser notadamente consistentes com as opiniões originais dos estudantes. Nesse sentido, os resultados provenientes das avaliações feitas pelos estudantes não podem ser descartados simplesmente como uma apreciação inadequada de jovens imaturos.

Outro aspecto importante referente aos possíveis vieses encontrados nas avaliações docentes é o momento de aplicação do instrumento. As avaliações conduzidas na metade do curso, a discussão dos resultados dessas avaliações com os estudantes podem servir de retroalimentação para o ensino. Não obstante, as avaliações conduzidas no final do curso são mais válidas como índice de efetividade do professor (Overall & Marsh, 1980). Por este motivo, a não ser que exista alguma razão especial, recomenda-se que a aplicação se dê ao final do curso.

A avaliação de desempenho, em Psicologia, pode ser verificada de diversas maneiras e por diferentes pessoas que trabalham diretamente com o avaliado. Na presente pesquisa a maneira proposta de se avaliar o desempenho de docentes foi a partir de avaliações realizadas pelos discentes. Nesse sentido, estudar as propriedades psicométricas de instrumentos de avaliação utilizados nesse processo pode garantir inferências confiáveis

sobre a qualidade do docente e orientar melhor as tomadas de decisões institucionais que visem a uma crescente qualidade do ensino.

Os atributos de um bom professor devem ser avaliados por instrumentos válidos e confiáveis e permitir que melhores conceitos lhe sejam atribuídos visando alcançar os objetivos propostos nos planos de ensino e conceitos inferiores a docentes que não conseguem desenvolver adequadamente o ensino proposto. Neste âmbito, é de significativa importância que os dirigentes de instituições de ensino utilizem instrumentos com tais propriedades, para que possam construir um retrato fidedigno do desempenho acadêmico do corpo docente, e, ainda, nortear a construção de ações que visem ao aperfeiçoamento dos professores. Também para o professor é importante que as informações contidas nos relatórios que recebem sejam válidas e fidedignas e que indiquem pontos relevantes a serem considerados em sala de aula para alcançar um nível elevado da qualidade de ensino. Diante disso, existem argumentos relevantes que justificam o desenvolvimento e aprofundamento deste estudo e a aplicação das técnicas estatísticas que permitem verificar a qualidade do instrumento de medida utilizado, dada sua significativa importância para o contexto educacional.

O presente trabalho considera as ações avaliativas do corpo docente como parte de um processo mais amplo de avaliação de instituições de ensino superior e as preocupações que envolvem esse processo. Considera, também, os princípios que devem nortear a avaliação institucional, a nova sistemática de avaliação do ensino superior brasileiro, a avaliação de desempenho docente como parte da avaliação institucional e de curso, e os cuidados com o processo e com os instrumentos utilizados.

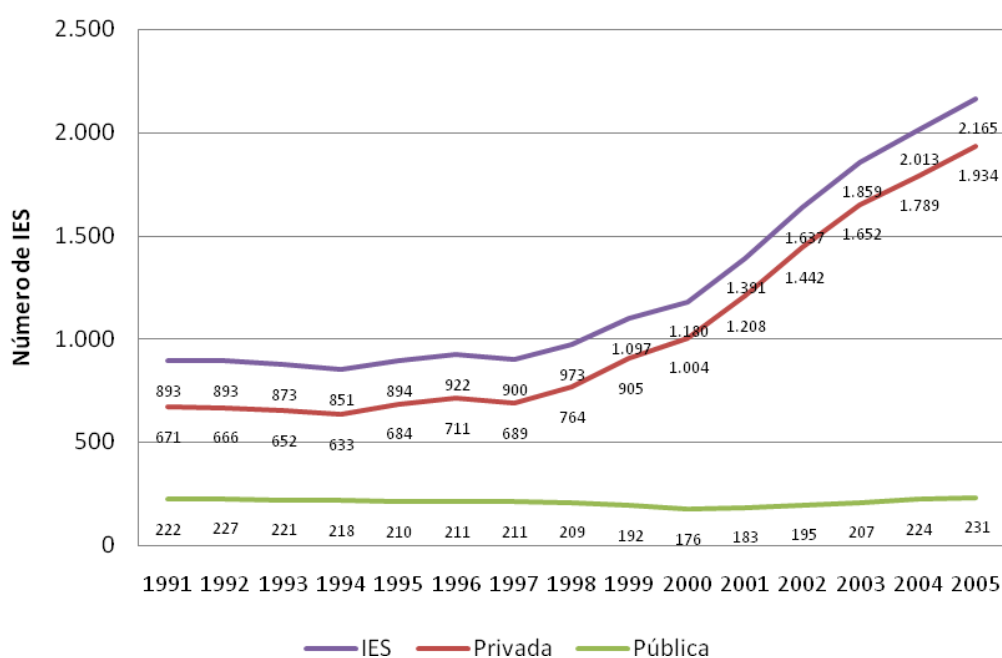
Com a finalidade de pesquisar sobre a avaliação do desempenho docente pelos estudantes de ensino superior, buscou-se por pesquisas referentes a este tema que tragam

resultados sobre a influência de diferentes variáveis e sobre a validade de instrumentos de medida utilizados nesse processo de avaliação. Essas pesquisas estão descritas nos capítulos seguintes.

SISTEMAS DE AVALIAÇÃO EDUCACIONAL E DE DESEMPENHO DOCENTE

O sistema de avaliação de uma instituição educacional deve seguir certos princípios de avaliação institucional tais como os recomendados por Cashin (1996), são eles: clareza das metas institucionais, propósito da informação a ser coletada, definição de suas fontes e de sua aceitabilidade técnica, envolvimento dos participantes, apoio dos níveis administrativos superiores, flexibilidade e legalidade, responsabilidades do corpo docente e de outros membros da instituição, treinamento dos avaliadores e supervisores para fornecerem o *feedback*, manutenção da confidencialidade, recompensa do desempenho efetivo e revisão periódica do sistema.

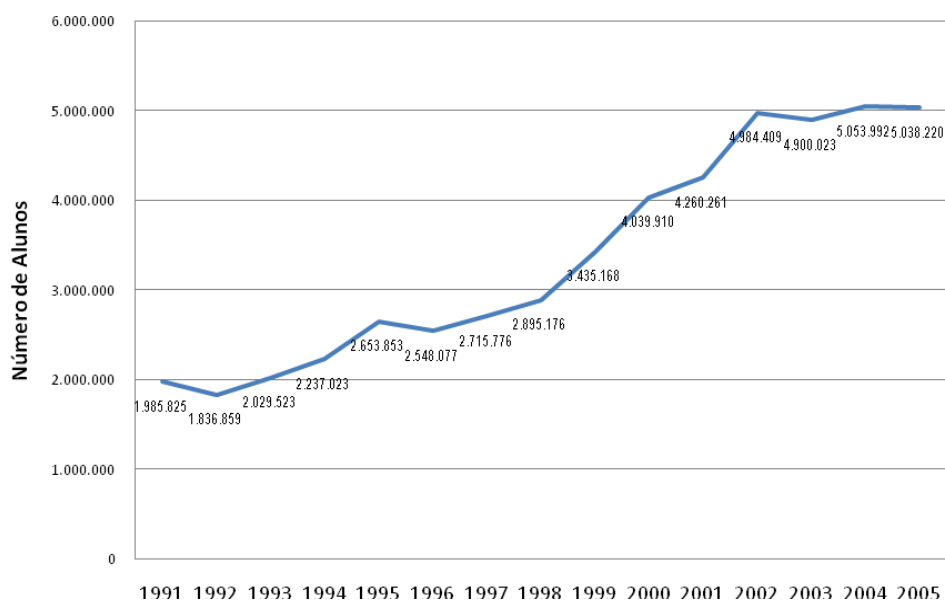
No Brasil, o Decreto nº 5.773, de 9 de maio de 2006 (MEC, 2006), dispõe sobre o exercício das funções de regulação, supervisão e avaliação de instituições de educação superior e cursos superiores de graduação e sequenciais no sistema federal de ensino. De acordo com as informações disponibilizadas no Sistema de Estatísticas Educacionais, em 2005, o Brasil detém 2.165 instituições, das quais 1.934 são mantidas pela iniciativa privada e 231 são públicas (MEC, 2007), conforme ilustrado na figura a seguir.



Fonte: Sistema de Estatísticas Educacionais (MEC, 2007)

Figura 1. Distribuição do número de IES do Brasil no período de 1991 a 2005 por categoria administrativa

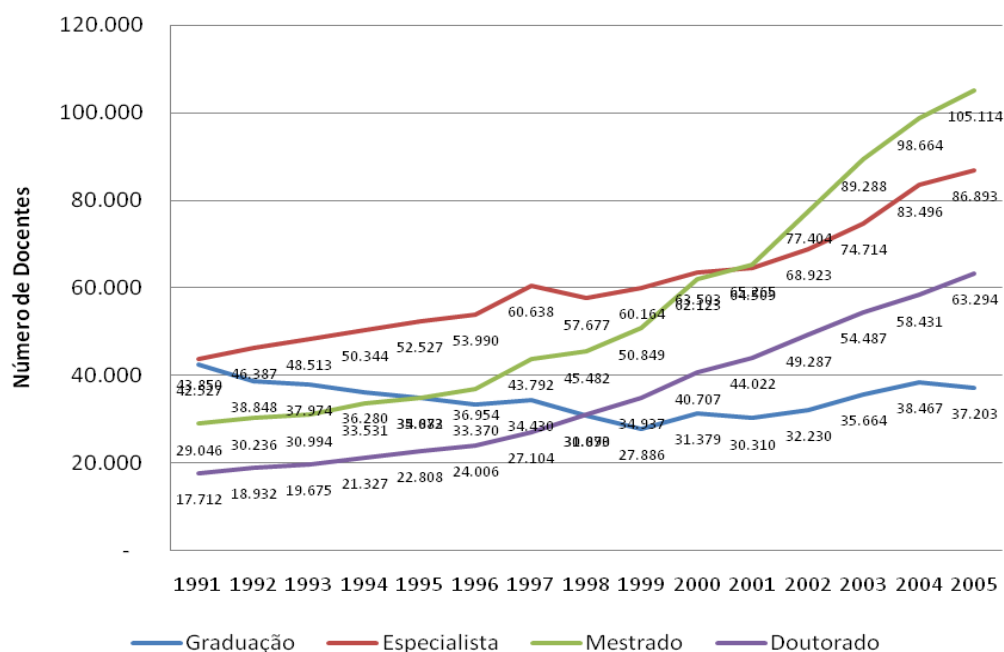
Percebe-se claramente o crescimento no número de IES, representando um aumento de 242% em relação a 1991. Esse aumento se deve basicamente ao número de IES privadas criadas a partir de 1997. Com a abertura de novas instituições, em 2005 o Brasil ofertou 2.117.449 vagas nas instituições privadas e, como consequência favorecendo o ingresso das pessoas nas IES representando, portanto, um crescimento de 370% em relação aos índices registrados em 1991. As maiores ofertas estão concentradas no turno noturno, compreendendo 1.475.120 vagas. A maior oferta de vagas deve-se a um acréscimo significativo do número de estudantes inscritos em IES, como mostra a Figura 2, só em 2005 o MEC registrou 5.038.220 estudantes inscritos em IES do Brasil (MEC, 2007).



Fonte: Sistema de Estatísticas Educacionais (MEC, 2007)

Figura 2. Distribuição do número de alunos inscritos nas IES do Brasil no período de 1991 a 2005

Os dados apresentados na Figura 2 mostram que o número de estudantes de ensino superior cresceu mais de 153% em relação a 1991, demandando um número cada vez maior de profissionais de ensino. No ano de 2005, 292.504 profissionais exerceram funções docentes, conforme a Figura 3, na qual se percebe o aumento significativo desses profissionais acompanhando a evolução do número de IES no Brasil já apresentada na Figura 1.



Fonte: Sistema de Estatísticas Educacionais (MEC, 2007)

Figura 3. Distribuição do número de docentes em exercício no período de 1991 a 2005 por titulação

Fica evidente na figura anterior a preocupação dos professores com o aprimoramento de sua qualificação que, é mais acentuada para o número de mestres. Decorrente do aumento significativo de pessoas envolvidas no processo educacional torna-se necessário o acompanhamento desse sistema no sentido de regular, supervisionar e avaliar as IES e a qualidade da educação que elas oferecem a seus estudantes. A fim de promover o acompanhamento e a melhoria da qualidade nas IES, a CONAES instituiu o Sistema Nacional de Avaliação da Educação Superior – SINAES, o qual constitui referencial básico para os processos de regulação e supervisão da educação superior, sendo competência do Ministério da Educação, por intermédio de suas secretarias, exercerem as funções de regulação e supervisão da educação superior em suas respectivas áreas de atuação (INEP, 2004).

Recentemente, de acordo com a assessoria de comunicação social (ACS) do MEC (2006), foram implementados ajustes no SINAES, de modo a realizar uma avaliação mais completa para o controle de qualidade dos cursos superiores. O atual sistema leva em conta: o processo de ensino, o processo de aprendizagem, a capacidade institucional e a responsabilidade do curso com a sociedade em geral. Cada um desses itens constitui um indicador parcial que, combinados, compõem o Índice do Desenvolvimento do Ensino Superior (IDES). Para a construção do indicador de ensino, passaram a ser utilizados dados do corpo docente de cursos de graduação relativos ao número e formação, dedicação às atividades docentes, publicações e outros trabalhos, além de avaliações de docentes feitas por alunos, quando houver (MEC, 2005). Assim, a avaliação institucional, tanto no que se refere à auto-avaliação quanto à avaliação externa, constitui-se em condição básica para o necessário aprimoramento do planejamento e gestão da instituição, uma vez que propicia a constante reorientação de suas ações (MEC, 2005).

No contexto do SINAES, a avaliação das instituições de educação superior deve abranger duas modalidades: a *avaliação interna*, a auto-avaliação, coordenada pela Comissão Própria de Avaliação (CPA), instituída obrigatoriamente pela instituição de ensino e constituída por representantes do corpo docente, do corpo discente e do pessoal técnico-administrativo, além de representantes da sociedade de cada instituição e orientada pelas diretrizes e pelo roteiro da auto-avaliação institucional da CONAES; e, *avaliação externa*, realizada por comissões designadas pelo INEP, tendo como referência os padrões de qualidade para a educação superior, expressos nos instrumentos de avaliação e nos relatórios da auto-avaliação.

A avaliação externa fundamentada no SINAES e pautada no projeto de auto-avaliação da instituição e em seus relatórios, bem como em instrumentos comuns a todas as

IES, tem por objetivo contribuir para o autoconhecimento e aperfeiçoamento das atividades desenvolvidas pela instituição, levando em consideração o importante princípio do respeito à diversidade e à identidade institucional. Nesse sentido, a auto-avaliação e a avaliação externa devem estar completamente articuladas, pois ambos os processos são concebidos como subsídios fundamentais para a formulação de diretrizes para as políticas públicas de educação superior e para a gestão das instituições, visando à melhoria da qualidade de suas ações.

As instituições que, por qualquer motivo, deixarem de atender às exigências dos parâmetros de avaliação adotados, sofrerão as sanções previstas na legislação, podendo até serem descredenciadas pelo MEC, mediante indicação da CONAES. Esse sistema de avaliação do ensino superior respeita os critérios estabelecidos pela CPA. O indicador de ensino leva em conta os diferentes mecanismos da avaliação do corpo docente, incluindo as avaliações realizadas pelos estudantes, quando houver. Todo este processo é acompanhado por uma comissão especializada determinada pela Comissão Técnica de Acompanhamento da Avaliação (CTAA).

Segundo o MEC (2005), desenvolver um processo de avaliação institucional é assumir como postulados, além da democracia institucional, da liberdade nas ações e ética no fazer, da articulação dialógica entre qualidade e quantidade e da sensibilidade institucional para mudança, os seguintes cinco princípios norteadores:

1. *globalidade*, avaliação de todos os elementos que compõem a instituição de ensino;
2. *comparabilidade*, busca de uma padronização de conceitos e indicadores;
3. *respeito à identidade das IES*, consideração de suas características;

4. *legitimidade*, adoção de metodologias e construção de indicadores capazes de conferir significado às informações, que devem ser fidedignas;
5. *reconhecimento*, por todos os agentes, da legitimidade do processo avaliativo, seus princípios norteadores e seus critérios.

A problemática da avaliação interna da instituição suscita preocupações quanto à construção dos instrumentos utilizados. O MEC reitera em seus princípios norteadores que os instrumentos a serem utilizados pelo INEP devem obedecer a princípios rígidos de consistência e validade. O mesmo se aplica aos instrumentos internos utilizados pelas CPAs. Como em qualquer teste psicológico, um instrumento de avaliação de docentes deve obedecer a requisitos técnicos para que as decisões baseadas em seus resultados possam ser respaldadas em dados válidos, confiáveis e precisos. Esses requisitos referem-se a evidências de validade e precisão do instrumento.

Corcoran e Fisher (1987) chamam atenção para o fato de que validade e precisão de um instrumento devem ser revistas periodicamente, para que os dados sejam atualizados e se tornem aplicáveis à população atual. Eles também recomendam a escolha de instrumentos cujo coeficiente de precisão seja da ordem de 0,80, pelo menos, e cuja validade tenha sido estabelecida por diversos estudos e não apenas por um único. Além desses critérios, os autores recomendam a observação de mais três aspectos ao selecionar um instrumento de avaliação: a extensão do mesmo, a aceitabilidade social de seu conteúdo e a facilidade de compreensão dos itens propostos.

Em conformidade com as considerações de Corcoran e Fisher (1987), um instrumento de avaliação de desempenho docente, parte de um sistema amplo de auto-avaliação institucional, deve receber atenção especial quanto aos requisitos técnicos

envolvidos, sobretudo quando decisões importantes são tomadas com base nos resultados obtidos por esse instrumento.

No entanto, ainda as informações geradas nos relatórios decorrentes da avaliação docente, quando realizada pelos estudantes da IES, são muito polêmicas. Muitos acreditam que os estudantes somente terão condições de avaliar seus professores após estarem desempenhando suas atividades em áreas correlatas aos conteúdos desenvolvidos nas diferentes disciplinas durante o curso. Outros alegam que os alunos seriam influenciados por fatores, tais como: descompromisso com essa avaliação, interesses pessoais divergentes, vínculo com o professor avaliado, relações com as instâncias burocráticas, administrativas e financeiras da instituição, preocupação com as próprias notas ou conceitos nas disciplinas cursadas, entre outros (Primi & Moraes, 2005).

Segundo Primi e Moraes (2005), a falta de maturidade dos alunos, principalmente a dos iniciantes recém-chegados do ensino médio de escolas da rede pública de ensino, onde esta prática de retratar o desempenho acadêmico do professor não é desenvolvida, pode ocasionar dados imprecisos. Na literatura, a falta de maturidade e experiência não é confirmada nas pesquisas relatadas por Tejedor e Monteiro (1990). Todos esses argumentos são justos e legítimos, porém há ausência de estudos que evidenciem as propriedades psicométricas, confiabilidade e validade, fundamentais na construção de um instrumento de avaliação com número reduzido de questões que possibilite nortear possíveis intervenções objetivando a qualidade de ensino, e a atividade docente que ocupa um papel importante no processo educacional.

A atividade docente é concebida por Tejedor e colaboradores (1988) como uma atividade instrumental, dirigida a soluções de problemas mediante a aplicação de teorias e técnicas científicas. Segundo o autor, o professor é conhecido como um profissional que

desempenha uma tarefa (e que, portanto, requer uma formação) prioritariamente técnica, e seu êxito dependem da habilidade de relacionar-se com a sala e resolver os problemas inerentes a esse relacionamento. E ainda, a habilidade requerida refere-se à integração inteligente e criativa do conhecimento e da técnica (Yinger, 1986 citado por Tejedor e colaboradores, 1988).

A AVALIAÇÃO DE DESEMPENHO DOCENTE

A busca pela qualidade do ensino está relacionada diretamente com as ações político-pedagógicas que permeiam o ambiente acadêmico. Inserido neste aspecto, o corpo docente de uma IES atua com importante papel para alcançar a efetiva qualidade. Visando este objetivo, as instituições necessitam de um corpo de professores atualizados nas respectivas áreas de conhecimento, possuidores de habilidades e técnicas pedagógicas suficientes adequadas para conduzir o processo de ensino-aprendizagem. Ao considerar o professor universitário como um dos pressupostos relacionados ao processo de ensino, sabe-se que nem todos os profissionais que se propõem a ministrar determinada disciplina o fazem com proficiência. É preciso que o professor esteja continuamente preparado para o exercício da docência para atender, adequadamente, às exigências dos novos estudantes que semestralmente chegam às universidades e os anseios da sociedade. Ciente da importância do papel que exerce o professor, o acompanhamento do seu desempenho, por meio, de avaliações sistemáticas possibilita identificar os aspectos relacionados ao processo de ensino-aprendizagem que necessitem de intervenção com o objetivo de aprimorar a efetiva qualidade de ensino. Diante dessa perspectiva, Doyle (1977, citado por Tejedor & Monteiro, 1990) considera três paradigmas que inspiram as diferentes estratégias de avaliação da eficácia do desempenho docente, quais sejam: o paradigma processo–produto, o paradigma processos mediadores e o paradigma interativo. Borich (1986 citado por Such, 1999) acrescenta aos paradigmas anteriores o paradigma prognóstico–produto.

Esses paradigmas caracterizam-se, em síntese, pelos aspectos descritos a seguir. No paradigma *prognóstico–produto*, são considerados indicadores de competência do professor, baseada em suas características físicas e psicológicas, 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.

No paradigma *processo–produto*, são consideradas as relações entre a conduta do professor (processo) e os resultados alcançados pelos alunos (produto). Nesse paradigma o professor é a causa imediata do aprendizado do estudante, consideradas as variáveis: ordem expositiva, sequência lógica, relacionamento entre temas, elaboração de sínteses, avaliação do aprendizado, uso de recursos de apoio pedagógico e relacionamento com os estudantes.

No paradigma *dos processos mediadores* o professor influencia certos processos que constituem as atividades cognitivas dos alunos, que darão posteriormente um lugar para a aprendizagem. A investigação deste paradigma se dá em laboratórios de estudos psicológicos. Ele compreende todos os aspectos do paradigma anterior, reconhece e utiliza uma variável adicional relativa ao contexto escolar e social do estudante, 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, conhecimentos técnicos do professor sobre aprendizagem, nível de conhecimento dos discentes, consciência metacognitiva e outras. Esse paradigma sustenta a hipótese de que o professor não influi de modo direto no aprendizado do estudante, e sim em alguns processos que constituem a atividade cognitiva deste, ocorrendo, então, o seu aprendizado.

No paradigma *interativo* a eficácia do professor em relação ao aprendizado dos discentes ocorre em contexto específico. Este paradigma compõe aspectos naturalistas e etnográficos, com objetivos fundamentalmente compreensivos e descritivos, visualiza e reconhece a complexidade das relações e trocas que ocorrem durante os processos 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 estudantes; uso de estratégias diversas para a solução de problemas, avaliação centrada na compreensão e do domínio do tema. Para uma melhor visualização das diferenças dos aspectos que caracterizam os quatro paradigmas descritos anteriormente, constitui-se a Figura 4, que representa graficamente as diferentes estratégias de avaliação docente segundo Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999).

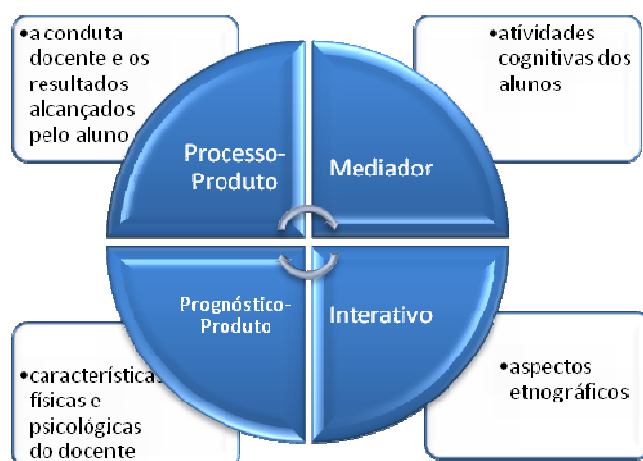


Figura 4. Paradigmas que inspiram as estratégias de avaliação da eficácia do desempenho docente segundo Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999).

Segundo Tejedor e colaboradores (1988), não existe acordo entre os pesquisadores sobre quais variáveis compreendidas no paradigma processo–produto relacionam-se melhor com o aprendizado do estudante. Além disso, a avaliação de professores será deficiente se forem utilizadas como único critério de referência o desempenho acadêmico dos estudantes. É necessário, portanto, obter medidas indiretas da competência docente, gerando um

modelo multidimensional de avaliação, que integre, quando possível, as dimensões específicas a todos esses paradigmas.

A seguir, estão apresentadas de forma esquemática as pesquisas que segundo Tejedor e Monteiro (1990) compreendem aos aspectos mais importantes que devem ser considerados na avaliação do professor universitário.

Braskamp, Brandenburg & Cry (1994, citado por Tejedor & Monteiro, 1990), adotaram o esquema prognóstico–processo–produto ilustrado abaixo.

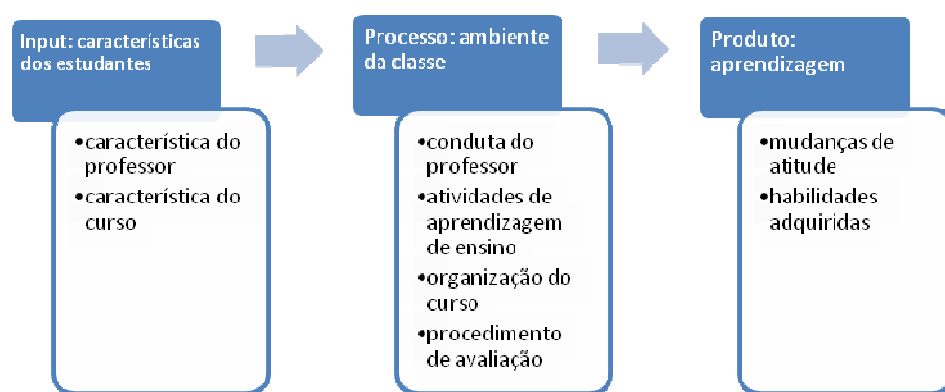


Figura 5. Esquema prognóstico-processo-produto adotado por Braskamp, Brandenburg & Cry (1994, citado por Tejedor & Monteiro, 1990)

Limitada a atividade especificamente docente, sem referência aos trabalhos de pesquisa e gestão, a proposta de Wilman (1988, citado por Tejedor & Monteiro, 1990) estabelece abaixo as condutas dos professores, que é o objeto prioritário da avaliação.

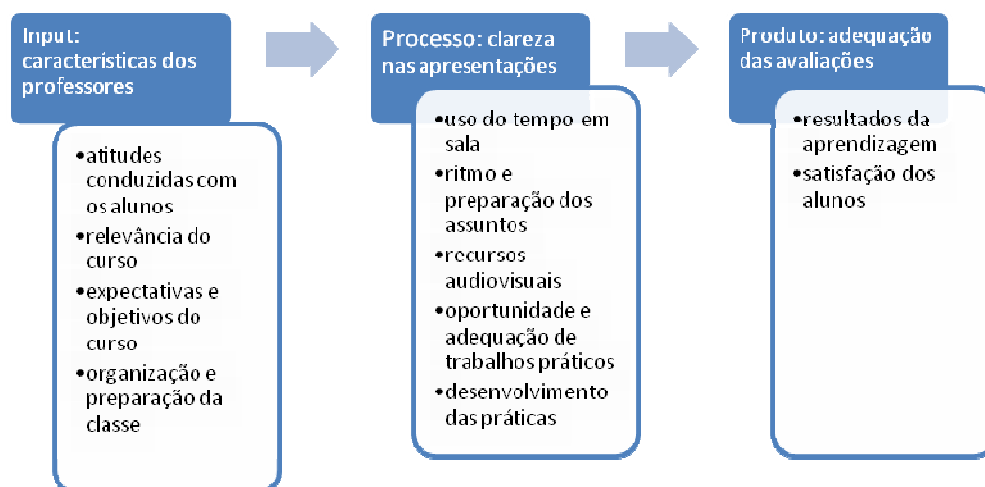


Figura 6. Esquema proposto por Wilman (1988, citado por Tejedor & Monteiro, 1990) sobre as condutas de um professor

Nessa mesma ótica, o trabalho apresentado por Rodríguez Espinar (1988, citado por Tejedor & Monteiro, 1990) sugere que os programas de avaliação da Universidade de Barcelona adotassem as referências técnicas e assistenciais necessárias aos seguintes indicadores por sua direta relação com a qualidade docente.

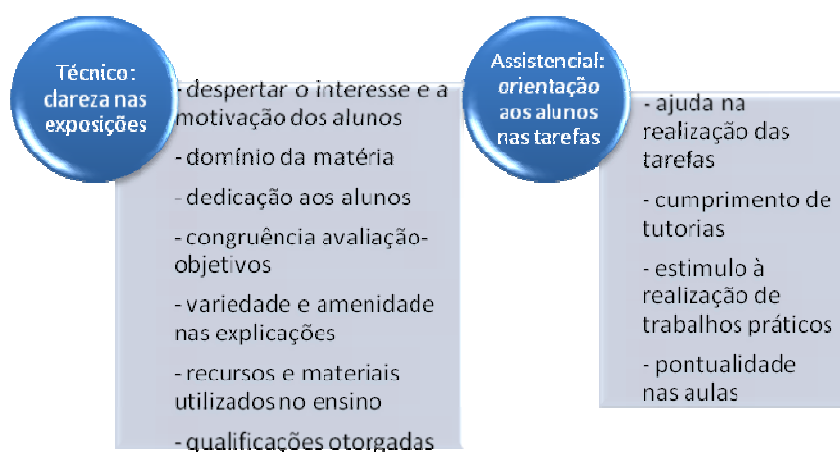


Figura 7. Referenciais técnicos e assistenciais identificados por Rodríguez Espinar (1988, citado por Tejedor & Monteiro, 1990) necessários na relação direta com a qualidade docente

A avaliação dos professores pelos alunos se orienta, sobretudo, a estimar o nível de qualidade de ensino universitário, a fim de contribuir progressivamente para sua melhora, e o método mais utilizado para coletar a opinião dos estudantes em relação ao ensino que recebe é o questionário onde se pede que os estudantes avaliem seus professores nos traços considerados mais relevantes para o ensino (Tejedor e colaboradores, 1988). Segundo o autor, resultados de pesquisas apontam os traços abaixo como os relacionados com aspectos da atividade docente, especialmente as inspiradas no paradigma processo–produto.

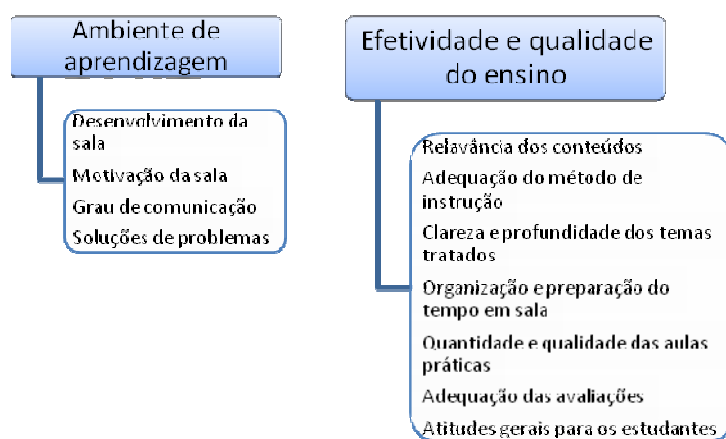


Figura 8. Representação gráfica dos traços relacionados com aspectos da atividade docente segundo Tejedor e colaboradores (1988)

O instrumento, objeto de estudo desta pesquisa, foi inspirado nos paradigmas de Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999). Os aspectos que compõem cada um dos itens contemplam os aspectos considerados em cada paradigma. A partir dos resultados de pesquisa empírica realizada por Primi e Moraes (2005) esses aspectos foram agrupados em quatro dimensões de qualidade docente.

Embora, seja de grande importância para a veracidade das informações verificar as propriedades psicométricas nos instrumentos de avaliação, é sabido que muitas vezes as instituições de ensino aceitam que seus instrumentos avaliativos encontram-se adequados sem mesmo terem realizado os devidos estudos comprobatórios.

O número de publicações que tratam desse tema na literatura brasileira ainda é reduzido, fator que favorece a existência de muitas questões ainda não respondidas sobre a validade de instrumentos construídos para avaliar o desempenho do professor universitário com base na opinião do estudante.

A carência de estudos, principalmente no Brasil, das propriedades psicométricas que compõem um instrumento avaliativo torna as questões de confiabilidade e de validade ainda pouco discutidas. Em razão dessa dificuldade, as instituições de ensino superior aceitam a hipótese de que seus instrumentos avaliativos são adequados, válidos, e passam a aplicá-los e tomar decisões a partir dos resultados obtidos sem os devidos estudos preliminares.

Os questionários definidos por Yaremko, Harari, Harrison e Lynn (1986, p. 186) como “um conjunto de perguntas sobre um determinado tópico que não testa a habilidade do respondente, mas mede sua opinião, seus interesses, aspectos de personalidade e informação biográfica” devem ser válidos e fidedignos. Portanto, é importante que tenham propriedades psicométricas que garantam a confiabilidade das inferências feitas com base nos resultados de sua aplicação.

O conceito de validade como o grau em que um teste realmente mede aquilo que ele se propõe a medir é definido por Anastasi e Urbina (2000) de forma objetiva. A precisão ou fidedignidade de um teste, segundo as autoras, refere-se à consistência dos escores obtidos

pelas mesmas pessoas em ocasiões diferentes; um traço medido por um teste deve se manter constante independentemente da ocasião (Pasquali, 2001; Anastasi e Urbina, 2000).

Outras definições podem ser encontradas na literatura, como a de Cronbach (1996) que utiliza três termos para definir validade: *validade de critério*, quando se compara o escore de uma medida psicológica com um critério externo, que pode ser preditiva ou concorrente; *validade de conteúdo*, quando os itens de um instrumento de medida abrangem o conteúdo considerado; *validade de construto*, quando os itens de um instrumento de medida estão de acordo com o referencial teórico empregado. Segundo as associações *American Educational Research Association*, *American Psychological Association* e *National Council on Measurement in Education* (1999) ao se verificar empiricamente a validade de um teste se encontram evidências de sua validade. Essas evidências são baseadas no conteúdo, no processo de respostas, na estrutura interna, nas relações com outras variáveis, evidência convergente e discriminante, generalização de validade ou nas consequências da testagem.

Fundamentalmente, todos os procedimentos para determinar a validade de um teste, segundo Anastasi e Urbina (2000), são voltados para as relações entre o desempenho e outros fatores independentes e observáveis relativos às características do comportamento considerado na análise. De acordo com as autoras, a *validade de construto* ocorre quando as evidências demonstram que a medida está de acordo com o construto teórico ou traço. As técnicas aplicadas para verificar o construto podem ser a correlação de teste com outros testes, a análise fatorial, a consistência interna, a validação convergente e discriminante. Como dito anteriormente, a validade e fidedignidade são aspectos cruciais na elaboração de qualquer instrumento, pois evidenciam que as medidas adotadas no processo avaliativo estão isentas das diversas fontes de erro que atuam durante o processo de medição. Cashin

(1995) acrescenta que os assuntos básicos em torno do instrumento de avaliação da docência pelos estudantes estão relacionados com a sua dimensionalidade, precisão e validade do instrumento. A dimensionalidade está associada ao número de características de um objeto (Hair e colaboradores, 2005). No caso de instrumento de avaliação dos professores, a dimensionalidade refere-se ao número de dimensões do instrumento de avaliação relevantes para a efetividade do ensino.

Quando se realizam medições repetidas das características em análise a confiabilidade segundo Malhotra (2001) garante o grau que uma escala produz resultados consistentes. Na presente análise, a confiabilidade das avaliações realizadas pelos estudantes é determinada pelo grau de consistência e uniformidade das respostas dos diferentes estudantes de uma mesma sala ao atribuírem pontos às características que indicam o seu nível de satisfação com os docentes. Portanto, a validade de um instrumento de medida pode ser considerada como o grau com que os indicadores unidimensionais de uma escala medem, com consistência interna, as distintas dimensões do instrumento que foram designados para medir (Ramos, 1986).

As avaliações dos estudantes como uma medida da qualidade do ensino recebido, de acordo com Marsh (1984), são difíceis de serem validadas, em virtude da existência de múltiplos critérios que definem um ensino eficaz. De acordo com Churchill (1979), uma medida tem validade, quando as diferenças entre os valores observados refletem diferenças verdadeiras sobre as características que se pretende medir e não sobre outros fatores. Neste sentido, a validade de uma escala de desempenho docente pode ser definida em razão das diferenças entre pontuações da escala, o que explica as verdadeiras características que distinguem um bom professor de outro, e não pela influência de outras variáveis.

Diversas pesquisas têm sido realizadas, conforme capítulo a seguir, que discutem diferentes aspectos relacionados à avaliação de desempenho do corpo docente, que auxiliaram na condução desta pesquisa.

PESQUISAS REFERENTES À AVALIAÇÃO DE DESEMPENHO DOCENTE

A seguir são relatados vários estudos a respeito da avaliação do desempenho docente pelos estudantes e a influência de variáveis como gênero, atribuições de notas pelo professor, interação no relacionamento professor–aluno, influência da personalidade do professor na avaliação, satisfação dos alunos com o desempenho de seus docentes, entre outras variáveis. São apresentadas, também, pesquisas referentes à dimensionalidade dos instrumentos de avaliação e as diferentes dimensões que cada instrumento apresenta.

Com o objetivo de verificar a percepção dos estudantes sobre a qualidade do ensino dos professores no âmbito da sala de aula, Pasquali (1984) aplicou um questionário composto por 92 itens a uma amostra de 354 estudantes de cursos superiores. Com o uso da técnica de análise fatorial esse instrumento foi reduzido a 71 itens distribuídos em sete dimensões que explicavam aproximadamente 49% da variância total. As dimensões foram intituladas como: o grande educador, desrespeito, laboratório, mestre, compreensão da disciplina, respeito, e imparcialidade.

Em outra pesquisa Blanco e Paz (1988) elaboraram um instrumento de avaliação composto de 27 itens, o qual foi aplicado aos 13.846 alunos da Universidade de Cantabria. Esse estudo resultou em três dimensões, que explicam 80% da variabilidade total, são eles: assistência à classe/organização, interação e domínio/entusiasmo do professor. A dimensão dois, denominada “interação”, manteve-se em todas as análises fatoriais, enquanto as demais foram interpretadas com significados distintos.

Neste mesmo ano, Owlia e Aspinwall (1988) encontraram um número reduzido de dimensões para avaliar o desempenho docente. Estes pesquisadores avaliaram a

multidimensionalidade de instrumento composto de 19 itens, destinados a medir a qualidade do ensino de Engenharia. A análise empírica resultou em quatro dimensões, que explicaram 61% da variabilidade total da escala. As dimensões foram rotuladas como: atitudes do professor, conteúdo, recursos acadêmicos e competência.

Ainda em 1988, uma estrutura fatorial, composta de oito dimensões, foi identificada por Tejedor e colaboradores (1988) ao avaliar a qualidade da docência por meio de um questionário composto de 49 itens, respondido por 44.323 estudantes da Universidade de Santiago. Os dados foram analisados com o emprego da técnica estatística de análise fatorial ortogonal e oblíqua com oito dimensões que explicaram 67% da variância total da escala. As dimensões encontradas foram: clareza, organização, e preocupação com os trabalhos; interação com os alunos; avaliação adequada; práticas e material de apoio; desenvolvimento do programa; dedicação, disciplina e plano de estudo; e, atualização informática.

Em anos posteriores, Casanueva, Periañez e Rufino (1997) estudaram, por meio da análise fatorial exploratória, a multidimensionalidade do instrumento destinado a avaliar a qualidade do serviço docente prestado pelo professor universitário, validado por uma amostra de 358 discentes cursando a mesma disciplina. Os resultados explicam 56% da variância total distribuída em cinco dimensões: atitudes pessoais do professor, competência expositiva, aspectos objetivos da preparação da aula, avaliação/formalidade e demanda de esforço.

Entre os estudos que avaliam a influência do gênero na avaliação docente Basow e Silberg (1987) analisam, por meio de dois instrumentos, 16 docentes de ambos os sexos de uma faculdade particular no interior dos Estados Unidos, realizada por aproximadamente 1000 estudantes, também de ambos os sexos. O primeiro instrumento foi o *Teacher Rating*

Form, composto por 26 itens do tipo *Likert* de 5 pontos, variando do valor 1, bem acima da média, ao valor 5, bem abaixo da média. A primeira questão tratava da habilidade geral de ensino, sua escala variando do valor 1, notável, a 5, pobre e inadequada. As demais questões eram distribuídas em 5 fatores: escolaridade, organização/clareza, interação instrutor-grupo, interação instrutor-aluno individualmente, dinamismo/entusiasmo. O segundo instrumento empregado foi a forma reduzida do *Bem Sex-Role Inventory* (BSRI), utilizado para avaliar como os alunos percebiam os traços masculinos ou femininos de seus professores. Os resultados sugerem cuidado ao levar em conta as avaliações de docentes por parte de estudantes, pois foi observado um efeito do gênero na avaliação dos traços masculinos ou femininos, pois a profissão de professor, em geral, tem sido percebida pelos estudantes como uma ocupação feminina. Por meio da técnica de *path analysis*, Erdle, Murray e Rushton (1985) verificaram a validade da avaliação de docentes pelos discentes por meio da relação entre as dimensões de personalidade, usadas como preditoras, e do comportamento instrucional, mediante avaliações feitas pelos alunos em relação aos professores. Os autores concluíram que a personalidade retratada no comportamento em sala de aula influencia na efetividade do ensino. Professores com alto nível de empatia tendem a ter comportamentos carismáticos, favorecendo o processo ensino-aprendizagem, em função do bom relacionamento mantido com a turma, ocasionando, por sua vez, avaliações mais positivas por parte dos alunos. Para os autores do estudo, o comportamento em classe é o mediador da relação entre personalidade e efetividade de ensino do professor, sendo que os alunos avaliam seus professores em relação aos comportamentos que eles apresentam na classe durante as aulas.

A tabela a seguir relaciona as estruturas fatoriais apontadas pelos autores com os aspectos avaliados no instrumento da presente pesquisa.

Tabela 1. Distribuição dos fatores de qualidade docente encontradas por diferentes autores segundo as dimensões de qualidade docente da presente pesquisa

Autor	Dimensões de qualidade docente da presente pesquisa				Outras dimensões
	Didática do Professor	Conteúdo	Processo de Avaliação	Relacionamento	
Pasquali (1984)	O grande educador; Mestre.	Compreensão da disciplina.		Respeito; Imparcialidade; Desrespeito.	Laboratório.
Basow e Silberg (1987)	Organização/ clareza; Dinamismo/ entusiasmo.			Interação instrutor-grupo; Interação instrutor-aluno; Individualmente.	Escolaridade
Blanco e Paz (1988)	Assistência à classe/ organização; Entusiasmo.	Domínio.		Interação	
Owlia e Aspinwall (1988)	Competência.	Conteúdo.		Atitudes do professor.	Recursos acadêmicos
Tejedor e colaboradores (1988)	Clareza, organização, e preocupação com os trabalhos. Dedicção.	Desenvolvimento do programa; Plano de estudo;	Avaliação adequada.	Interação com os alunos; Disciplina	Práticas e material de apoio; Atualização em informática.
Casanueva, Periañez e Rufino (1997)	Competência expositiva; Aspectos objetivos da preparação da aula; Demanda de esforço.		Avaliação/ Formalidade.	Atitudes pessoais do professor.	Demanda de esforço.

Percebe-se na Tabela 1 que os itens que compõe os instrumentos de pesquisa dos diferentes autores foram considerados para o presente estudo. Além disso, os indicadores de qualidade docente foram inspirados nos paradigmas apontados por Tejedor e colaboradores (1988) como relevantes para a atividade docente. Este fato sustenta a relevância de se manter as quatro dimensões apresentadas na Tabela 1, como indicadores de qualidade do desempenho docente pelos estudantes.

Para Castro (1992), os discentes reconhecem a dedicação dos docentes, a sua motivação para o ensino, a excelência da didática utilizada, e outras. Todas essas características são importantes para alcançar um ensino competente. Portanto, parecem 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). Os autores analisaram os níveis de satisfação dos alunos quanto ao desempenho de seus professores em sala de aula após o término de 100 disciplinas. Um ano depois da conclusão do curso, os mesmos alunos foram convidados a expor o nível de satisfação com os mesmos professores. Dessa pesquisa, resultou um coeficiente de estabilidade médio de 0,83, que levou os autores a concluir que os alunos conservam as mesmas opiniões sobre a qualidade do ensino dos professores ao longo do tempo.

Em uma revisão histórica sobre a validade e utilidade da avaliação dos estudantes em relação à qualidade de ensino, Greenwald (1997) descreve alguns estudos e preocupações de pesquisadores relativos a este tema. Alunos com notas baixas tendiam a avaliar negativamente o professor, enquanto os com notas altas tendiam a fazer o oposto. Em 1980, os estudos correlacionais, de modo geral, davam suporte à avaliação de docentes por estudantes, considerando que as eventuais interferências poderiam ser atribuídas a outros fatores, não necessariamente relacionados às notas escolares recebidas pelos

estudantes (efetividade do ensino, motivação dos estudantes, interesse prévio do estudante pelo assunto abordado, entre outros).

Os resultados oriundos das avaliações dos docentes realizadas pelos estudantes ainda são muito questionados pelo corpo de professores das IES. Alguns alegam que tais resultados são inconsistentes em razão da falta de maturidade dos alunos, principalmente das séries iniciais, para perceber que determinadas condutas dos professores fazem-se necessárias para favorecer o aprendizado do aluno. Outros alegam que somente professores de áreas correlatas e com uma maior titulação acadêmica possam fazê-lo. Outros, ainda, questionam que os resultados das avaliações podem ser influenciados por variáveis como o número de alunos em sala, o conteúdo programático de determinadas disciplinas e outras, que não são controladas durante o processo avaliativo. De acordo com Greenwald (1997), as principais discussões sobre este assunto surgem a partir de 1990: se avaliações respondidas pelos alunos seriam influenciadas por variáveis que não estariam relacionadas diretamente ao ensino efetivo, a unidimensionalidade ou multidimensionalidade da medida, e se as avaliações realizadas estariam sendo devidamente empregadas para a melhoria do ensino.

Esse propósito é a principal preocupação por parte dos participantes envolvidos no processo, corpo docente, discente e diretivo. A avaliação empregada de forma errada gerará informações que não são condizentes com seu propósito. Armstrong (1998) ressalta temer que uma avaliação docente mal conduzida, ao invés de propiciar uma melhoria da qualidade do ensino, acabe provocando o efeito contrário. A inquietação do autor sobre essa possibilidade é que tal avaliação produza outros efeitos negativos, levando professores a desinteressar-se em contribuir para o processo de aprendizagem do aluno. Buck (1998) apontou que existe o risco dos cursos superiores tornarem-se divertidos e envolventes,

porém vazios de conteúdo, efetivamente incapazes, nesse sentido, de preparar os alunos para o mercado de trabalho.

Greenwald e Gillmore (1998) afirmaram que o problema acima está na dúvida se o aluno atribui um conceito favorável somente aos comportamentos carismáticos do professor não contemplando os aspectos trabalhados pelo docente para seu aprendizado. Se essa afirmação for verdadeira, esse tipo de avaliação não seria útil para a melhoria da qualidade do ensino. Porém, sua prática continua em uso por falta de melhor alternativa e por ser um indicador na avaliação externa realizada pelo Ministério da Educação nas IES do Brasil..

Atualmente, existem diversos autores que ressaltam os pontos positivos e negativos da avaliação docente realizada pelos alunos como citam Gillmore e Greenwald (1999). Segundo esses autores, é importante destacar que alguns erros apresentados por esse tipo de avaliação podem ser corrigidos por meio estatístico, removendo os efeitos das interferências identificáveis. Uma das formas sugeridas de proceder essa correção é introduzindo-se um sistema em que os escores médios são ajustados por meio de análise de regressão múltipla, que permite determinar a porcentagem da variância que não deveria ser interpretada como medida da qualidade instrucional. Estes autores, ainda, reconhecem a possibilidade do estabelecimento de um círculo vicioso, como consequência de um sistema de avaliação mal planejado e mal utilizado, no qual as queixas dos alunos causam uma revisão nos conteúdos, excluindo muitas vezes os tópicos mais difíceis, o que leva indiretamente a um aumento nas notas acadêmicas, que, por sua vez, podem contribuir para um aumento na pontuação atribuída pelo aluno na avaliação do docente.

Losak (1994) adverte que os administradores de faculdades e universidades devem evitar enfocar somente testes e números, esconder os problemas e utilizar o auto-estudo como um instrumento promocional da instituição. Outro aspecto polêmico é a divulgação

dos resultados das avaliações; nesse sentido, Howell e Symbaluk (2001) investigaram as opiniões de alunos e professores de uma faculdade canadense sobre a divulgação das avaliações feitas por alunos sobre o ensino oferecido pela escola. Em conclusão, de modo geral, os alunos mostraram-se favoráveis a essa divulgação, listando uma série de vantagens, entre elas o acesso a informações sobre a instituição, que facilitaria a escolha de cursos e de professores. Em contrapartida, os professores, apesar de concordarem com esses dois argumentos, mostraram-se desfavoráveis à divulgação dessas informações, preocupados com sua privacidade e com o efeito que uma avaliação negativa pode ter, mesmo que o professor já tenha modificado suas estratégias de ensino. Os autores sugerem como alternativa para essa divergência, a criação de páginas institucionais na internet, em que informações sobre cursos e professores sejam divulgadas sem comprometer a privacidade dos envolvidos. Além disso, sugerem tornar mais acessíveis as informações sobre os cursos antes da escolha pelos alunos ou publicar as diretrizes dos cursos antes do início da seleção, no começo de cada semestre ou ano.

Não existem conclusões precisas sobre os fatores decorrentes do processo de divulgação das avaliações docente no Brasil, no entanto, mesmo assim, as instituições de ensino acreditam que devam divulgar os resultados de suas avaliações à comunidade interna, temendo cair no descrédito caso não o façam. Mesmo assim, elas tomam alguns cuidados, como por exemplo, preservar o nome dos participantes envolvidos, divulgar somente resultados amplos, além de restringir os relatórios à comunidade interna.

Diante desse cenário, este estudo pretende oferecer uma contribuição sobre as evidências da validade de um instrumento de avaliação docente em uma IES. Pretende-se verificar a fidedignidade e validade de construto desse instrumento. As características avaliadas do docente referem-se a características de qualidade dos professores refletidas

nos resultados positivos provenientes da aplicação do instrumento e que não podem ser diretamente e perfeitamente medidas.

Assim, o presente estudo pretende avaliar as propriedades psicométricas do instrumento de avaliação docente concebido com quatro indicadores que, em conjunto, podem constituir um único fator: didática, conteúdo, avaliação e relacionamento.

OBJETIVO GERAL

Este trabalho tem como objetivo verificar evidências de validade de construto de um instrumento de avaliação docente realizada por estudantes de uma universidade particular do estado de São Paulo.

Objetivos específicos

- Determinar a consistência interna do instrumento de avaliação.
- Examinar a estrutura fatorial do instrumento, utilizando-se das técnicas de análise fatorial exploratória.

MÉTODO

Fonte de dados

O presente estudo foi realizado tomando por base um banco de dados que contém informações de 698 docentes avaliados pelos seus alunos, 59% do gênero masculino, 25% doutores, 36% mestres, 28% especialistas e 11% graduados. No total, 8.644 alunos avaliaram voluntariamente, via internet, em ambiente seguro, seus professores em 32 cursos das 1.341 disciplinas de diferentes áreas de conhecimento ministradas em quatro câmpus universitários nos semestres de 2005 e 2006. Esses resultados estão apresentados nas tabelas a seguir.

Tabela 2. Distribuição do número de docentes avaliados por ano, semestre e gênero

Ano	Semestre	Feminino		Masculino		Total de docentes avaliados
		Nº	%	Nº	%	
2005	1º	243	41%	348	59%	591
	2º	235	45%	290	55%	525
2006	1º	245	43%	329	57%	574
	2º	239	42%	329	58%	568
Total de docentes avaliados		288	41%	410	59%	698

Observa-se que o maior número de docentes avaliados é do gênero masculino, isto se deve ao fato de que há um maior percentual de docentes do gênero masculino na instituição. Esta proporção se mantém em todos os semestres.

Quanto à titulação docente (Tabela 3) a maior parte dos docentes é pelo menos mestre, o que revela a preocupação da instituição em contratar e manter docentes com maior titulação, o que pode contribuir para o ensino de qualidade.

Tabela 3. Distribuição do número de docentes avaliados por titulação, ano e semestre

Titulação acadêmica	Ano/Semestre				Total de docentes avaliados
	2005		2006		
	1º	2º	1º	2º	
Doutores	160	138	157	144	175
Mestres	220	207	207	206	253
Especialistas	155	132	156	163	195
Graduados	29	26	27	28	75
Total de docentes avaliados	591	525	574	568	698

Nota: A maior parte dos docentes foi avaliada nos quatros semestre.

Na instituição, um total de 698 docentes atuou no período de 2005 a 2006, sendo que a maior parte foi avaliada por seus estudantes em pelo menos um dos semestres letivos. Desses docentes, a maior parte é de mestres ou doutores (61%), revelando a preocupação da instituição com a qualificação de seus docentes.

Conforme dados apresentados na Tabela 4, aproximadamente 50% dos estudantes participaram das avaliações docentes em 2005 e 2006.

Tabela 4. Distribuição do número de estudantes que responderam a avaliação docente por ano, semestre e gênero dos estudantes

Ano	Semestre	Feminino		Masculino		Total de estudantes
		Nº	%	Nº	%	
2005	1º	2.776	58%	1.984	42%	4.760
	2º	1.534	55%	1.279	45%	2.813
2006	1º	2.271	58%	1.628	42%	3.899
	2º	2.628	57%	2.002	43%	4.630
Total de estudantes		4.945	57%	3.699	43%	8.644

Dos 8.644 estudantes que avaliaram os docentes, a porcentagem média do gênero feminino é de 57% variando de 55% a 58% do 1º semestre de 2005 ao 2º semestre de 2006.

Tabela 5. Distribuição do número de estudantes que responderam a avaliação docente por curso, ano e semestre

Curso	Ano/Semestre				Número de estudantes participantes	
	2005		2006			
	1º	2º	1º	2 º		
Administração	706	454	594	805	1.403	52%
Análise de Sistemas	27	22	9	6	33	50%
Arquitetura e Urbanismo	112	61	72	85	164	46%
Ciência da Computação	53	54	57	64	115	48%
Ciências	15	9	3	1	22	79%
Ciências Biológicas	65	33	50	58	113	53%
Direito	1.275	548	891	1.103	2.129	52%
Educação Física	53	25	44	42	80	44%
Enfermagem	116	75	125	131	242	47%
Engenharia Ambiental	20	13	23	29	45	50%
Engenharia Civil	43	28	45	47	98	55%
Engenharia de Computação	121	85	121	102	208	47%
Engenharia de Produção	15	13	25	22	45	57%
Engenharia Elétrica	111	98	113	113	204	47%
Engenharia Industrial - Modalidade Mecânica	37	32	35	40	81	52%
Engenharia Mecânica - Automação e Sistemas	218	188	176	243	435	52%
Farmácia	221	171	296	338	557	50%
Fisioterapia	170	91	93	88	236	51%
Hotelaria	84	54	45	43	109	45%
Letras	78	34	66	66	119	48%
Matemática	80	31	39	50	109	53%
Medicina	141	94	106	154	257	45%
Nutrição	38	24	60	65	94	46%
Odontologia	119	46	102	110	230	56%
Pedagogia	203	110	161	172	327	49%
Psicologia	464	316	412	461	841	48%
Química	40	24	25	21	60	50%
Química Industrial			17	43	46	73%
Serviço Social	83	58	86	84	169	50%
Sistemas de Informação	15	9	9	7	22	52%
Técnicas em Gestão de Sistemas Informatizados				32	33	100%
Turismo	64	44	32	37	79	43%

Nota: Dos 8.644 alunos 61 estão vinculados a mais de um curso

Os cursos que representam a maior participação no processo avaliativo foi Direito (24%) e Administração (16%), fato que se deve ao elevado número de alunos matriculados nestes cursos em relação aos demais.

Material

O instrumento de avaliação, objeto deste estudo (Anexo 1), foi elaborado a partir de um instrumento de avaliação de desempenho docente composto por 15 indicadores, apoiado-se nos paradigmas apresentados por Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999).

O instrumento original foi aplicado no período de 30 de outubro a 13 de dezembro de 2004. Os professores foram avaliados em cada um dos 15 itens abaixo pelo aluno.

1. A clareza das explicações do professor nos assuntos tratados;
2. A relação que o professor estabelece entre os aspectos conceituais da disciplina e as aplicações práticas;
3. A integração que o professor faz desta disciplina com as outras do curso (interdisciplinaridade);
4. O conhecimento que o professor tem dos conteúdos desta disciplina;
5. A organização das aulas práticas;
6. A atenção dispensada à orientação/supervisão de Estágios, Prática Jurídica, TCC, ou Relatório Final;
7. O aproveitamento do tempo destinado às aulas;
8. O cumprimento do programa apresentado;

9. A adequação dos métodos de ensino utilizados pelo professor;
10. A adequação das avaliações em relação aos conteúdos trabalhados e discutidos na disciplina;
11. A clareza/objetividade dos critérios de avaliação usados pelo professor;
12. A análise que o professor faz dos resultados das avaliações visando à melhoria da aprendizagem;
13. A motivação/incentivo dado pelo professor para você participar das aulas;
14. A atenção do professor para você responder suas dúvidas;
15. A postura ética do professor (respeito aos alunos, professores, à profissão e à instituição).

Estudos realizados por Primi e Moraes (2005) indicam que, embora o desempenho docente seja um construto multidimensional, segundo Greenwald (1997), o instrumento composto de 15 indicadores pode ser reduzido, sem perder informação, a um instrumento composto de indicadores abrangentes que medem uma opinião geral que o aluno forma do desempenho docente. Estão apresentados na tabela 5 à avaliação realizada por 3.689 estudantes avaliando 568 professores em 36 cursos. A análise fatorial realizada pelos autores indicam que 66% da variância total é explicada pelo primeiro componente da estrutura fatorial e que a precisão do instrumento é alta, com um coeficiente alfa de Cronbrach igual a 0,965.

As estatísticas descritivas, as cargas fatoriais e comunalidade de cada um dos 15 itens do instrumento de avaliação docente aplicado em 2004 estão apresentados a seguir.

Tabela 6. Estatísticas descritivas dos itens da escala de avaliação docente de 15 itens - 2004

Item	Média	Desvio Padrão	N*	Cargas Fatoriais	Comunalidade
A clareza das explicações do professor nos assuntos tratados	4,0	1,089	25.267	0,862	0,743
A relação que o professor estabelece entre os aspectos conceituais da disciplina e as aplicações práticas	4,1	1,052	25.167	0,862	0,743
A integração que o professor faz desta disciplina com as outras do curso (interdisciplinaridade)	3,9	1,062	24.720	0,819	0,671
O conhecimento que o professor tem dos conteúdos desta disciplina	4,5	0,842	25.285	0,752	0,565
A organização das aulas práticas	4,0	1,111	22.858	0,845	0,715
A atenção dispensada à orientação/supervisão de Estágios, Prática Jurídica, TCC, ou Relatório Final	4,0	1,078	16.655	0,805	0,648
O aproveitamento do tempo destinado às aulas	4,1	1,063	24.541	0,816	0,665
O cumprimento do programa apresentado	4,2	0,964	23.997	0,809	0,654
A adequação dos métodos de ensino utilizados pelo professor	4,0	1,106	23.982	0,875	0,766
A adequação das avaliações em relação aos conteúdos trabalhados e discutidos na disciplina	4,1	1,044	23.039	0,810	0,656
A clareza/objetividade dos critérios de avaliação usados pelo professor	4,1	1,056	22.806	0,811	0,658
A análise que o professor faz dos resultados das avaliações visando melhoria da aprendizagem	4,0	1,061	21.741	0,816	0,666
A motivação/incentivo dado pelo professor para você participar das aulas	3,9	1,201	22.926	0,831	0,690
A atenção do professor para você responder suas dúvidas	4,3	1,005	22.989	0,789	0,622
A postura ética do professor (respeito aos alunos, professores, à profissão e a instituição)	4,5	0,885	22.910	0,680	0,463

* Número de avaliações feitas pelos estudantes aos respectivos professores em suas disciplinas.

Fonte: Primi e Moraes (2005)

Neste instrumento as maiores médias foram atribuídas aos indicadores que tratam sobre: “O conhecimento que o professor tem dos conteúdos desta disciplina”, alcançando 4,5 de satisfação do alunado, e referente “à postura ética do professor (respeito aos alunos, professores, à profissão e a instituição)”, praticamente obtendo o mesmo

desempenho. Os itens que obtiveram os piores resultados na avaliação do aluno foram: “A integração que o professor faz desta disciplina com as outras do curso (interdisciplinaridade)”, com média de 3,9 pontos; “A motivação/incentivo dado pelo professor para você participar das aulas”, alcançando a mesma pontuação.

De maneira geral a média do desempenho do corpo docente da instituição pode ser considerada alta, pois todos os itens estão próximos da pontuação 4, valor acima do ponto médio da escala.

Segundo Tejedor e Monteiro (1990) a decisão para a escolha do número de itens dependerá dos objetivos da avaliação: quando se pretende que o professor melhore seu trabalho individual, haverá de incluir um número de itens elevado; se o objetivo é tratar de obter dados conclusivos e não formativos, bastarão poucos itens de caráter geral. No estudo proposto levou-se em consideração um número reduzido de itens com o objetivo de maximizar a participação e o completo preenchimento do instrumento eletrônico em tempo razoável, favorecendo, assim, a qualidade da resposta, evitando erros do tipo de *halo*, geralmente encontrados em questionários longos. Segundo esses autores, as pesquisas que visam definir um bom docente, a partir da análise fatorial de um instrumento de avaliação docente, oscilam de quatro a dez dimensões. Neste estudo foram consideradas quatro dimensões: a didática, o conteúdo, a avaliação e o relacionamento.

Para que os estudantes avaliem seus docentes as quatro dimensões foram apresentadas e relacionadas ao estudante na seguinte forma de quatro perguntas: *didática* corresponde a clareza das explicações nos assuntos tratados, a adequação dos métodos de ensino, o aproveitamento do tempo destinado às aulas, o cumprimento do programa apresentado e a diversidade de recursos didáticos; *conteúdo* aborda o conhecimento que o professor tem dos conteúdos desta disciplina; *avaliação* corresponde a questões inerentes à

adequação das avaliações em relação aos conteúdos trabalhados, à objetividade dos critérios de avaliação, à análise e encaminhamentos que o professor faz dos resultados das avaliações visando a melhoria da aprendizagem; e *relacionamento*, que tange ao incentivo dado pelo professor para a participação nas aulas, à atenção do professor para responder às dúvidas, à postura ética do professor ante os alunos, colegas, profissão e Instituição.

O instrumento é uma escala do tipo *likert* de 5 pontos: péssimo (1), ruim (2), regular (3), bom (4) e excelente (5). Quando o aluno responde que o indicador não se aplica/não sei avaliar é atribuído o valor zero (adotado como *missing*).

Procedimento

Para a realização da pesquisa que gera os dados analisados neste estudo foi disponibilizado pela instituição o instrumento de avaliação docente (Anexo 1), a partir do primeiro semestre de 2005, a todos os alunos da graduação, em meio eletrônico, via internet, em página que só pode ser acessada pela identificação acadêmica e senha individual do aluno, independentemente do local em que ele esteja realizando o acesso.

Os alunos receberam um e-mail personalizado convidando-os a realizarem a avaliação. Os participantes foram sensibilizados a participarem da avaliação por mensagens disponibilizadas no portal da instituição, cartazes em salas de aulas e corredores, e aviso oral pelos representantes de classe.

Apesar da ampla divulgação, a participação do aluno é voluntária. Ao decidir participar, o aluno acessa o *link* específico na página do aluno correspondente à avaliação docente e passa a responder aos itens relativos ao professor que aparece identificado e que ministrou aulas naquele semestre para aquele aluno em determinada disciplina. Caso o professor ministre mais de uma disciplina para o mesmo aluno, ele será novamente

indicado para ser avaliado. Quando termina a avaliação de uma dimensão, aparece na tela do computador o nome dimensão avaliada, com o respectivo docente em suas disciplinas e, assim, sucessivamente, até que todos os professores e dimensões tenham sido avaliados.

Os passos que compõe o processo de avaliação docente estão ilustrados abaixo.

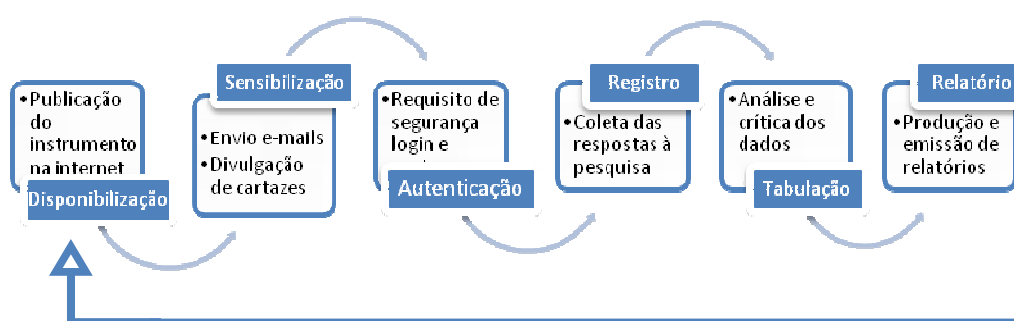


Figura 10. Fluxo do processo de avaliação docente

A aplicabilidade do instrumento se apoiou nos pressupostos da teoria de troca social recomendados por Dillman (1978, citado por Pasquali, 2003) com o objetivo de maximizar a participação dos alunos. O autor recomenda que para aumentar a participação dos respondentes, é necessário: minimizar o esforço do respondente para a participação na pesquisa, maximizar as recompensas e estabelecer um elo de confiança.

No âmbito operacional, as ações adotadas com base nas recomendações feitas por Dillman (1978, citado por Pasquali, 2003) foram:

- minimizar a dificuldade para a participação do respondente – o instrumento foi disponibilizado na internet, em linguagem adequada ao ambiente

acadêmico. Esse meio de coleta das informações possibilita ao aluno responder no momento, local e horário em que achar oportuno, podendo ser nos computadores disponibilizados pela instituição, em sua residência, no trabalho, em *lan-house*, *cibers-café*. Outra vantagem no emprego desta tecnologia no processo avaliativo está no fato de o sistema construir um questionário personalizado, levando em consideração as disciplinas específicas cursadas por cada aluno, indiferentemente se caracterizem como dependências ou como regulares. Possibilita, ainda, ao respondente a interrupção e o retorno ao questionário sem perda de informações.

- maximizar as recompensas – a maior recompensa entendida pelo aluno em responder ao instrumento está no fato de os resultados contribuírem para a melhoria da qualidade no ensino.
- estabelecer confiança que a recompensa será concedida – a legitimidade deste item está nas reuniões pedagógicas realizadas com o corpo docente e representantes do corpo discente para discussões e futuras ações decorrentes dos resultados obtidos na avaliação.

ANÁLISE DE DADOS

A consistência interna do instrumento é verificada através técnicas estatística que visem verificar a homogeneidade da amostra dos itens do teste. As técnicas mais utilizadas são: duas metades, a técnica de Kuder-Richardson e o alfa de Cronbach.

Na técnica de duas metades, o teste é dividido em duas partes equivalentes (metades), e a correlação é calculada nos escores obtidos nas duas metades. A técnica de Kuder-Richardson elaborada em 1937 baseia-se na análise individual de cada item do teste Cronbach em 1951 mostrou que esta técnica produz um coeficiente de precisão do teste que corresponde à média dos coeficientes de todas as metades em que o teste possa ser dividido. O alfa de Cronbrach é uma extensão da técnica de Kuder-Richardson.

As propriedades estatísticas de validade do instrumento de avaliação docente podem ser investigadas pelo uso das técnicas multivariadas, como a análise fatorial exploratória. A análise fatorial consiste em analisar os inter-relacionamentos entre as variáveis observadas. Ela pode ser empregada para examinar os padrões ou relações latentes para um grande número de variáveis e determinar se a informação pode ser condensada ou resumida em um conjunto menor de fatores ou componentes (Hair e colaboradores, 2005). Conforme descrito pelo autor, o objetivo da análise fatorial exploratória é resumir as inúmeras informações coletadas pelo instrumento de pesquisa em um conjunto menor de componentes ou fatores, ou seja, determinar uma estrutura subjacente aos dados que conserve o máximo possível da variabilidade contida nas variáveis analisadas.

Hair e colaboradores (2005) apresentam dois tipos de análise fatorial: análise fatorial do tipo R e do tipo Q. A análise do tipo R avalia as relações entre as variáveis para identificar quais grupos formam as dimensões latentes (fatores). A análise do tipo Q forma

grupo de respondentes ou casos com base em sua similaridade em um conjunto de características.

Pretende-se, com o emprego desta técnica multivariada de análise de dados, confirmar por meio das respostas obtidas dos alunos ao longo de quatro semestres letivos, a variabilidade dos níveis de satisfação em relação ao desempenho do conjunto de professores no âmbito da sala de aula. Outro aspecto relevante do estudo é verificar a estrutura dimensional do instrumento, para o que desenvolveremos a análise fatorial exploratória do tipo R.

Hair e colaboradores (2005) apresentam diferentes estratégias de rotação fatorial, que visam à manipulação ou ao ajuste dos eixos fatoriais para conseguir uma solução fatorial mais simples e pragmaticamente mais significativa. Elas podem se dividir em: as que mantêm a ortogonalidade dos fatores latentes – rotação ortogonal; e as que não mantêm a ortogonalidade dos fatores latentes – rotação oblíqua.

Existem diversas estratégias de rotação ortogonal as principais são *varimax*, *equimax* e *quartimax*. O método de rotação *varimax* é o mais utilizado no processo de validação de construto. O seu uso é justificado pelo fato de que o método busca a simplificação das colunas da matriz fatorial, ou seja, o método busca minimizar o número de indicadores com altas cargas num fator latente.

O propósito desta análise é encontrar uma estrutura simplificada que explique o máximo da variabilidade total das variáveis selecionadas. O método dos componentes principais é o adequado para a estimação dos fatores, pois considera a variância total e determina fatores que contêm pequenas proporções de variância única e, em alguns casos, variância do erro (Hair e colaboradores, 2005).

Para verificar o grau de correlação existente entre as variáveis de escala métrica (e a direção dessa correlação – se positiva ou negativa), utilizaremos o coeficiente de correlação de *Pearson*, também chamado de “coeficiente de correlação produto–momento” ou simplesmente de “r de Pearson”. Este coeficiente, representado pela letra “r”, assume apenas valores entre -1 e 1, onde: $r = 1$, significa uma correlação perfeita positiva entre as duas variáveis, $r = -1$, significa uma correlação negativa perfeita entre as duas variáveis – isto é, se uma aumenta, a outra sempre diminui e $r = 0$, significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir outra dependência que seja “não linear”. Assim, o resultado $r = 0$ deve ser investigado com o emprego das técnicas apropriadas.

Com relação ao número de sujeitos da amostra, Hair e colaboradores (2005) recomendam ter pelo menos 5 vezes mais observações do que o número de variáveis a serem analisadas. Para uma análise fatorial exploratória consistente as medidas de distribuição estão adequadas quando cumprem os critérios de linearidade, homocedasticidade e normalidade. Um modo de determinar essas características é o teste de esfericidade de *Bartlett*, um teste estatístico que verifica a hipótese de que todos os elementos da diagonal principal da matriz de correlação são iguais à unidade e os demais elementos sejam zero. O teste de *Bartlett* é sensível à normalidade presente na estrutura de dados, ele fornece a probabilidade estatística de que a matriz de correlação tenha correlações significantes entre pelo menos algumas das variáveis.

Outra medida para verificar o grau de intercorrelações entre as variáveis é o índice *Kaiser-Meyer-Olkin* (KMO). Ele compara a grandeza dos coeficientes de correlação observados com os coeficientes de correlações parciais para o conjunto de indicadores.

RESULTADOS E DISCUSSÕES

As propriedades psicométricas do instrumento de avaliação docente estão apresentadas a seguir. A tabela abaixo corresponde às estatísticas descritivas do instrumento de avaliação docente aplicado nos semestres de 2005 e 2006.

Tabela 7. Estatísticas descritivas da avaliação docente realizada pelo discente por ano, semestre, item do instrumento e câmpus

Ano	Semestre	Item	Câmpus												Total de estudantes		
			1			2			3			4					
			M	DP	Nº	M	DP	Nº	M	DP	Nº	M	DP	Nº	M	DP	Nº
2005	1º	Didática	4,1	1,0	15.885	4,0	1,1	7.468	4,1	1,0	5.87	3,9	1,0	1.581	4,1	1,0	30.811
		Conteúdo	4,5	0,8		4,4	0,9		4,5	0,8		4,4	0,8		4,5	0,8	
		Avaliação	4,1	1,0		4,1	1,0		4,1	1,0		4,1	0,9		4,1	1,0	
		Relacionamento	4,3	1,0		4,2	1,0		4,2	1,0		4,2	1,0		4,2	1,0	
	2º	Didática	4,1	1,1	8.367	4,0	1,1	4.854	4,1	1,1	2.544	3,9	1,1	1.418	4,0	1,1	17.183
		Conteúdo	4,4	0,9		4,3	1,0		4,4	1,0		4,3	0,9		4,4	0,9	
		Avaliação	4,1	1,0		4,0	1,0		4,1	1,0		3,9	1,0		4,0	1,0	
		Relacionamento	4,2	1,1		4,2	1,1		4,3	1,1		4,1	1,0		4,2	1,1	
2006	1º	Didática	4,1	1,0	11.788	4,0	1,1	6.100	4,1	1,1	4.187	3,9	1,1	2.024	4,1	1,1	24.099
		Conteúdo	4,5	0,8		4,4	0,9		4,4	0,9		4,3	0,9		4,4	0,9	
		Avaliação	4,1	1,0		4,0	1,0		4,1	1,0		4,0	1,0		4,1	1,0	
		Relacionamento	4,3	1,0		4,2	1,1		4,3	1,0		4,1	1,1		4,2	1,0	
	2º	Didática	4,1	1,0	14.148	4,0	1,1	7.204	4,2	1,0	5.217	4,0	1,1	2.077	4,1	1,0	28.646
		Conteúdo	4,4	0,9		4,4	0,9		4,5	0,8		4,3	0,9		4,4	0,9	
		Avaliação	4,1	1,0		4,1	1,0		4,2	1,0		4,0	1,0		4,1	1,0	
		Relacionamento	4,3	1,0		4,2	1,0		4,3	1,0		4,2	1,0		4,3	1,0	
Total		Didática	4,1	1,0	50.188	4,0	1,1	25.626	4,1	1,0	17.825	3,9	1,1	7.100	4,1	1,1	100.739
		Conteúdo	4,4	0,9		4,4	0,9		4,5	0,9		4,3	0,9		4,4	0,9	
		Avaliação	4,1	1,0		4,1	1,0		4,1	1,0		4,0	1,0		4,1	1,0	
		Relacionamento	4,3	1,0		4,2	1,1		4,3	1,0		4,2	1,0		4,2	1,0	

M = média; DP = desvio padrão

Conforme demonstrado acima, as médias obtidas pelo desempenho dos professores ao longo dos quatro semestres letivos permaneceram altas sem variabilidades significativas. No geral, a maior média alcançada corresponde ao indicador relacionado ao

conteúdo, o qual também evidenciou a menor variabilidade em relação aos demais indicadores ($M = 4,4$; $DP = 0,9$) destacando o câmpus 3 ($M = 4,5$; $DP = 0,9$). A menor média alcançada e a maior variabilidade foram obtidas pelos indicadores relacionados com a Didática ($M = 4,1$; $DP = 1,1$), seguidos pelo item Relacionamento ($M = 4,2$; $DP = 1,0$). O câmpus 4 apresentou menores médias nos itens: Didática ($M = 3,9$; $DP = 1,1$), Conteúdo ($M = 4,3$; $DP = 0,9$) e Relacionamento ($M = 4,1$; $DP = 1,0$) em comparação aos outros três.

Corroborando com os estudos realizados por Primi e Moraes (2005) embora o desempenho docente seja um construto multidimensional, segundo Greenwald (1997), o instrumento composto de 15 indicadores pode realmente ser reduzido, sem perder informação, fato que pode ser observado nas igualdades das médias gerais obtidas nos instrumentos de 2004 ($M = 4,1$, $DP = 1,0$), 2005 ($M = 4,2$, $DP = 1,0$) e 2006 ($M = 4,2$, $DP = 1,0$). Embora os autores Primi e Moraes (2005) expressem que não existe perda de informações, entende-se que a mesma refere-se a aspectos gerais representados pelas médias de cada dimensão e não nas características específicas dentro de cada dimensão. As médias gerais praticamente não se alteram nos quatro semestres ao utilizar o instrumento que contém as quatro dimensões (tabela 6).

A matriz de correlação abaixo representa as intercorrelações existentes entre o conjunto de variáveis investigadas. Nessa matriz, os valores localizados na diagonal principal serão iguais a 1, isso porque cada uma das variáveis se correlacionará perfeitamente (linearmente) consigo mesma. Os elementos das diagonais secundárias são as correlações entre todos os pares correspondentes da variável.

A próxima tabela descreve as correlações existente entre os itens, demonstrando que estas são significativas e com razoável linearidade entre as variáveis em estudo.

Tabela 8. Matriz de correlação da avaliação docente pelo discente realizada nos semestres de 2005 e 2006.

Item	Estatísticas	Didática	Conteúdo	Avaliação	Relacionamento
Didática	Correlação de Pearson	1	0,706*	0,710*	0,707*
	Significância		0,000	0,000	0,000
	N	100739	100739	100739	100739
Conteúdo	Correlação de Pearson	0,706*	1	0,622*	0,612*
	Significância	0,000		0,000	0,000
	N	100739	100739	100739	100739
Avaliação	Correlação de Pearson	0,710*	0,622*	1	0,664*
	Significância	0,000	0,000		0,000
	N	100739	100739	100739	100739
Relacionamento	Correlação de Pearson	0,707*	0,612*	0,664*	1
	Significância	0,000	0,000	0,000	
	N	100739	100739	100739	100739

* A correlação é significativa ao nível de 0,01.

A maior correlação encontrada foi entre os indicadores Didática e Conteúdo $r(100738) = 0,706, p < 0,01$. A menor correlação apresentada foi entre Relacionamento e Conteúdo $r(100738) = 0,612, p < 0,01$.

De acordo com as recomendações dos autores Hair e colaboradores (2005) e Pereira (1999), a amostra utilizada neste estudo é adequada para a análise fatorial ($KMO = 0,837$). O teste de esfericidade de *Bartlett* indicou que as correlações são estatisticamente diferentes de zero ($\chi^2(6, N=100739) = 229.898,78; p < 0,001$).

A comunalidade é a quantia total de variância que uma variável original compartilha com todas as outras variáveis incluídas nas análises.

Tabela 9. Comunalidade extraída da avaliação docente pelo discente realizada nos semestres e 2005 e de 2006

Item	Inicial	Extração
Didática	1,000	0,815
Conteúdo	1,000	0,714
Avaliação	1,000	0,745
Relacionamento	1,000	0,738

Nota: Método de Extração: Análise de Componentes Principais

O item Didática é a variável que apresenta à maior comunalidade, os demais apresentam os valores semelhantes. Abaixo apresentamos o total da variância explicada no instrumento de avaliação docente.

Tabela 10. Autovalor e variância explicada dos componentes da avaliação docente

Componente	Autovalor Inicial	Porcentagem da variância	
		% Absoluta	% Acumulada
1	3,0	75,3	75,3
2	0,4	10,0	85,3
3	0,3	8,4	93,7
4	0,3	6,3	100,0

Método de Extração: Análise de Componentes Principais

Nesta análise o instrumento apresentou um único componente extraído, respondendo a 75,3% do total da variância acumulada, o que impossibilita a utilização dos métodos de rotação.

O total da variância explicada pelo instrumento (75,3%) está acima das pesquisas encontradas na literatura: Tejedor e colaboradores (1988) com o instrumento

aplicado na Universidade de Santiago após o emprego da técnica estatística de análise fatorial ortogonal e oblíqua, resultando numa estrutura fatorial composta por oito dimensões, que explicando 67% da variância total da escala, o Questionário para Avaliação de Ensino (TFQ) aplicado em Hong Kong, desenvolvido um e analisado empiricamente por Shah e colaboradores (2002) evidenciou uma estrutura fatorial composta por quatro fatores correspondendo a 63% da variância total, o instrumento desenvolvido por Owlia e Aspinwall (1988) composto por 19 itens, destinados a medir a qualidade do ensino de Engenharia, na análise empírica da multidimensionalidade o instrumento resultou em quatro dimensões, que explicando 61% da variabilidade total da escala, o instrumento destinado a avaliar a qualidade do serviço docente prestado pelo professor universitário, validado por uma amostra de 358 discentes cursando a mesma disciplina aplicado por Casanueva, Periañez e Rufino (1997) explicando 56% da variância total. A variância total explicada pelo instrumento ficou acima dos resultados relatados por Pasquali (1984) na aplicação do instrumento no Distrito Federal, o qual foi composto por 92 itens a uma amostra de 354 estudantes de cursos superiores. Por meio da análise fatorial esse instrumento foi reduzido a 71 itens distribuídos em 7 dimensões que explicavam aproximadamente 49% da variância total. É importante ressaltar que também foram encontrados resultados na literatura obtiveram índices acima de 76,3% encontrado no instrumento como o questionário desenvolvido Blanco e Paz (1988) composto de 27 itens, o qual foi aplicado aos 13.846 alunos da Universidade de Cantabria. Esse estudo resultou em três dimensões, que explicam 80% da variabilidade total.

A extração de um componente pode ser ilustrada por meio do gráfico de autovalor, Isso denota a unidimensionalidade do instrumento.

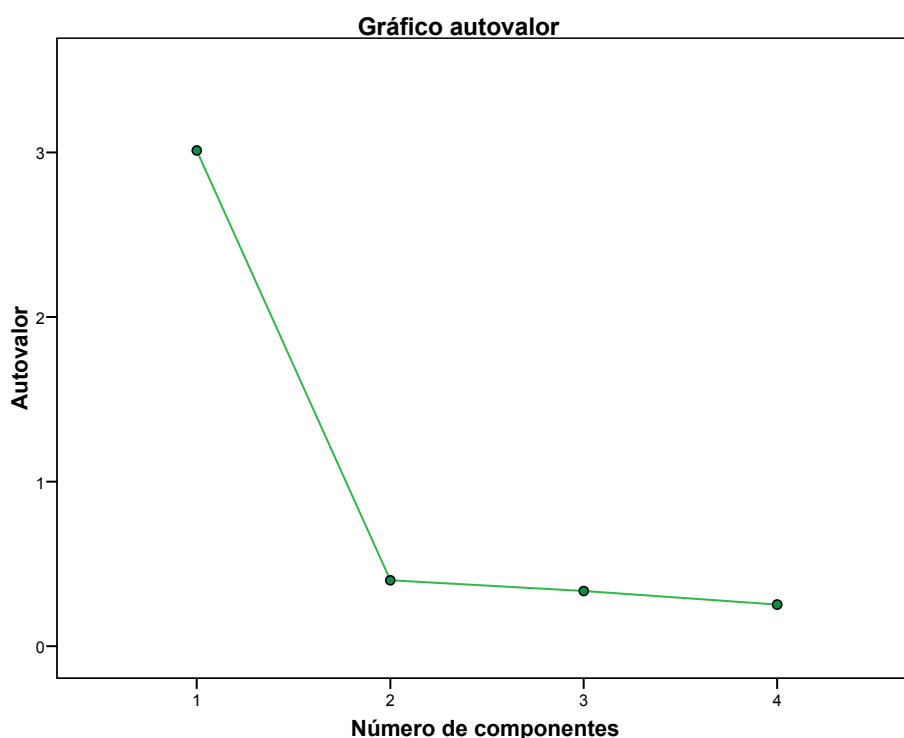


Figura 11. Gráfico autovalor

Os autores Tejedor e Monteiro (1990) e Lowman (2004) concordam que existe certo sentimento entre o corpo docente de que alunos com boas notas atribuem pontuações mais altas para os seus docentes do que alunos com notas mais baixas, ou que professores que não apresentam bom desempenho tentem “comprar” alunos atribuindo-lhes notas mais altas.

Estratificando a análise pela nota obtida pelo aluno menor ou igual a 4 (percentil 5) e maior ou igual a 9,5 (percentil 95), o resultado da variância acumulada em ambos os casos foi acima de 73% (79% para os alunos que obtiveram notas iguais ou abaixo de 4 e 73% para os alunos que alcançaram notas iguais ou acima de 9,5).

Abaixo distribuímos as médias alcançadas pelos docentes atribuídas pela nota do aluno, estratificada de acordo com o percentil 95 e 5.

Tabela 11. Distribuição das médias alcançadas pelos docentes atribuídas pelo aluno, estratificada de acordo com os percentis P₉₅ e P₅.

Ano	Semestre	Didática		Conteúdo		Avaliação		Relacionamento	
		Nota acima de 9,4	Nota abaixo de 4,1	Nota acima de 9,4	Nota abaixo de 4,1	Nota acima de 9,4	Nota abaixo de 4,1	Nota acima de 9,4	Nota abaixo de 4,1
2005	1º	4,3	3,7	4,6	4,2	4,3	3,7	4,5	3,9
	2º	4,3	3,7	4,5	4,1	4,3	3,7	4,5	3,8
2006	1º	4,3	3,6	4,6	4,1	4,4	3,6	4,5	3,7
	2º	4,3	3,7	4,5	4,1	4,3	3,6	4,5	3,8

Podemos perceber que os alunos que obtiveram notas inferiores a 4,1, tenderam a atribuir conceitos inferiores à avaliação dos professores em relação aos alunos que obtiveram um melhor desempenho acadêmico. Esses fatores podem influenciar na validação do instrumento, no entanto, segundo Lowman (2004) e Tejedor e Monteiro (1990), a literatura indica que não existe uma conclusão sobre as relações entre esses fatores e que a falta de acordo entre os resultados se deve a razões de ordem metodológica.

Na análise realizada pelo gênero no aluno em ambos os sexos o total da variância explicada foi de (75%). Podemos observar na tabela a seguir que não existem diferenças significativas nas avaliações realizadas pelos alunos e alunas de acordo com as médias distribuídas ao longo dos semestres, corroborando, assim, com os estudos realizados por Tejedor e Monteiro (1990) e Centra e Gaubatz (1999).

Tabela 12. Distribuição das médias alcançadas pelos docentes segundo o gênero do aluno

Ano	Semestre	Didática		Conteúdo		Avaliação		Relacionamento	
		Feminino	Masculino	Feminino	Masculino	Feminino	Masculino	Feminino	Masculino
2005	1º	4,1	4,0	4,5	4,4	4,1	4,0	4,3	4,2
	2º	4,1	3,9	4,4	4,3	4,1	3,9	4,3	4,1
2006	1º	4,1	4,0	4,5	4,4	4,1	4,0	4,3	4,2
	2º	4,1	4,1	4,5	4,4	4,1	4,0	4,3	4,2

Outras pesquisas citadas por (Tejedor & Monteiro, 1990 e Centra & Gaubatz, 1999) denotam que as alunas são mais críticas que os alunos, tendendo a valorar melhor as professoras que os professores, outras, ainda, indicam que as alunas valoram melhor os professores, em determinadas dimensões, que os alunos. A tabela abaixo apresenta a distribuição das médias das avaliações pelo gênero do aluno e professor.

Tabela 13. Distribuição das médias alcançadas pelos docentes de acordo com o gênero do docente e do aluno

Ano	Semestre e gênero do docente	Didática		Conteúdo		Avaliação		Relacionamento	
		Feminino	Masculino	Feminino	Masculino	Feminino	Masculino	Feminino	Masculino
2005	1º Feminino	4,1	4,0	4,5	4,4	4,1	4,0	4,3	4,2
	Masculino	4,1	4,1	4,5	4,5	4,1	4,0	4,3	4,2
	2º Feminino	4,1	3,8	4,4	4,1	4,1	3,8	4,3	4,0
	Masculino	4,1	4,0	4,4	4,3	4,1	4,0	4,3	4,2
2006	1º Feminino	4,1	3,9	4,4	4,3	4,1	4,0	4,2	4,2
	Masculino	4,2	4,1	4,5	4,4	4,2	4,1	4,3	4,2
	2º Feminino	4,1	4,0	4,4	4,3	4,1	4,0	4,3	4,2
	Masculino	4,1	4,1	4,5	4,4	4,1	4,1	4,3	4,3

Conforme observamos as médias apresentadas na tabela não correspondem aos resultados relatados na literatura, mas conforme ressaltado pelos autores pode-se perceber que tais resultados são difíceis de interpretar e apresentam resultados muitas vezes

contraditórios, pois a avaliação pode sofrer influência do papel social que, em ambos os sexos, é influenciado pela cultura.

A confiabilidade da dimensão foi medida pelo coeficiente alfa de Cronbach. Para o total de professores avaliados em suas respectivas disciplinas no geral foi 0,89; ao longo dos quatro semestres analisados o coeficiente se manteve acima de 0,884.

Tabela 13. Coeficiente de fidedignidade alfa de Cronbach da escala por ano e semestre

Ano	Semestre	Alfa de Cronbach
2005	1 °	0,884
	2 °	0,897
2006	1 °	0,891
	2 °	0,890
Total		0,890

Para cada câmpus o coeficiente também apresentou um padrão semelhante. Overall e Marsh (1980) analisaram os níveis de satisfação dos alunos quanto ao desempenho de seus professores em sala de aula ao final de cem disciplinas. Um ano após a conclusão do curso, os mesmos alunos foram convidados a expor o nível de satisfação com os mesmos professores. Dessa pesquisa, resultou um coeficiente de estabilidade médio de 0,83, o que se aproxima do coeficiente encontrado na avaliação. O coeficiente se mantém estável quando estratificado por câmpus, permanecendo acima de 0,882.

Ao se considerar que a confiabilidade do indicador deve superar 0,5 (Hair e colaboradores, 2005), os resultados demonstrados na tabela indicam que todos os indicadores permanecem consistentes em todos os câmpus e semestres avaliados. Esses resultados confirmam que as quatro dimensões que compõem o instrumento de avaliação

são consistentes e que não diferem quando estratificamos pelo turno em que o curso é ministrado: diurno (0,888), matutino (0,884), noturno (0,890), e vespertino (0,895). Pelo gênero do aluno: masculino (0,889) e feminino (0,890). De modo semelhante, os valores permanecem iguais, quando comparamos os alunos que obtiveram nota igual ou abaixo de 4,0 (0,908) com os que alcançaram notas superiores ou iguais a 9,5 (0,878). Quando estratificado por curso, todos os coeficientes ficaram acima de 0,837.

CONCLUSÃO

Esta pesquisa teve como objetivo buscar evidências de validade de construto para confirmar as dimensões ou fatores latentes subjacentes ao instrumento de avaliação docente. O questionário de avaliação docente realizada pelos discentes de uma universidade particular do interior do Estado de São Paulo, utilizado nesta pesquisa, engloba um conjunto de quatro dimensões que, *a priori*, definem um “bom docente”, à luz dos paradigmas apresentados por Doyle (1977, citado por Tejedor & Monteiro, 1990) e Borich (1986 citado por Such, 1999).

O questionário utilizado nesta pesquisa parte do pressuposto, indicado por Primi e Moraes (2005), que o instrumento de avaliação docente composto inicialmente por 15 itens, e aplicado nessa mesma instituição de ensino em 2004, poderia ter seu número de questões reduzidas. Para a comprovação da validade do instrumento reduzido composto de quatro dimensões, foi utilizada a base de dados proveniente do processo de avaliação docente ocorrido nos primeiros e segundos semestres letivos de 2005 e 2006.

Procurando atender a esse objetivo empregou-se a técnica multivariada de análise de dados conhecida como análise fatorial exploratória para verificar a validade de construto, isto é se o instrumento proposto é unidimensional. Com a aplicação desta técnica pode-se concluir que o instrumento é unidimensional, e que 75,3% da variância total é explicada pelo primeiro componente da estrutura fatorial, valor considerado alto. Além disso, o instrumento mostrou-se consistente internamente, com um coeficiente alfa de Cronbach igual a 0,890.

Os resultados indicam que não houve mudanças significativas nas médias gerais do desempenho do corpo docente em sala de aula, e que não há necessidade de se utilizar

formulários extensos, com inúmeras perguntas, para alcançar os indicadores globais de qualidade docente. Vale ressaltar que quanto mais indicadores se utiliza em um instrumento de avaliação maior é a possibilidade de se identificar fatores específicos de qualidade docente. No presente estudo o objetivo principal era identificar indicadores globais de qualidade docente e não de características mais detalhadas do desempenho do docente.

Do ponto de vista das limitações desta pesquisa, é preciso ressaltar que, outros estudos devem ser realizados considerando o efeito do tamanho das salas de aula, da dificuldade da disciplina, da situação socioeconômica e de outros aspectos que pudessem interferir nos resultados da avaliação docente. Além disso, recomenda-se que sejam realizadas pesquisas junto ao corpo docente e discente para investigar o uso dos resultados das avaliações dos professores como instrumento norteador de melhoria da qualidade de ensino.

BIBLIOGRAFIA

- American Educational Research Association, American Psychology Association & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. New York: American Educational Research Association.
- Almeida, A. J. (2002). O professor e a valorização de sua atividade docente. In: R. C. Feltran, *Avaliação na educação superior* (pp. 19-160). Campinas: Papirus.
- Anastasi, A., & Urbina, S. (2000). *Testagem Psicológica* (7 ed.). (M. A. Veronese, Trad.) Porto Alegre, RS, Brasil: Artmed.
- Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 79(3), 308-314.
- Blanco, L. S., & Paz, J. J. (1988). Evaluacion de la docencia mediante cuestionario de alumnos: Universidade de Cantabria (curso 86-87). *Studia Pedagogica*, 20, 41-71.
- Buck, D. (1998). Student evaluations of teaching measure the intervention, not the effect. *American Psychologist*, 53(11), 1224-1225.
- Casanueva, C., Periañez, R., & Rufino, J. E. (1997). Calidad percibida por el alumno en el servicio docente universitario: Desarrollo de una escala de medida. (M. Ruiz, Ed.) *XI Congreso Nacional AEDEM Lleida*, 2, 27-34.
- Cashin, W. E. (Setembro de 1995). *Student ratings of teaching: the research revisited*. Acesso em 28 de Dezembro de 2006, disponível em Education Resources Information Center: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/d2/44.pdf

- Cashin, W. E. (Janeiro de 1996). *Developing an effective faculty evaluation system*. Acesso em 28 de Dezembro de 2006, disponível em Education Resources Information Center: <http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED395536>
- Castro, C. M. (1992). E quem avalia os professores? Dois pontos: teoria e prática em educação, 2(13), 12-13.
- Centra, J. A., & Gaubatz, N. B. (1999). *Is there gender bias in student evaluation of teaching?* New Jersey: Educational Testing Service.
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, (16), 64-74.
- Corcoran, K., & Fisher, J. (1987). *Measure for clinical practice: a sourcebook*. New York, U.S.A.: Free Press.
- Cronbach, J. (1996). *Fundamentos da testagem psicológica* (5 ed.). Porto Alegre, RS, Brasil: Artes Médicas.
- Cunha, M. I. (1999). *O bom professor e sua prática* (9 ed.). Campinas, SP, Brasil: Papirus.
- Erdle, S., Murray, H. G., & Rushton, J. P. (1985). Personality, classroom behavior, and student ratings of college teaching effectiveness: a path analysis. *Journal of Educational Psychology*, 77(4), 394-407.
- Fuentes-Medina, M. E., & Sánchez, J. R. (1999). Evaluación docente: Hacia una fundamentación de la auto evaluación. *Revista electrónica interuniversitaria de formación del profesorado*, 2(1), 353-368.
- Gillmore, G. M., & Greenwald, G. M. (1999). Using statistical adjustment to reduce biases in student ratings. *American Psychologist* (54), 518-519.

- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52(11), 1182-1186.
- Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11), 1209-1217.
- Greenwald, A. G., & Gillmore, G. M. (1998). How useful are student ratings? Reactions to comments on the Current Issues Section. *American Psychologist*, 53(11), 1228-1229.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2005). *Análise Multivariada de Dados* (9 ed., Vol. I). (M. A. Gouvêa, Ed., A. S. Sannt'Anna, & A. Chaves, Trans.) Porto Alegre, RS, Brasil: Artmed Editora S.A.
- Howell, A. J., & Symbaluk, D. G. (2001). Published student ratings of instruction: revealing and reconciling the views of students and faculty. *Journal of Educational Psychology*, 93(4), 790-796.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (Setembro de 2004). *Sistema Nacional de Avaliação e Progresso da Educação Superior - Uma nova sistemática de avaliação do Ensino Superior Brasileiro*. Acesso em 28 de Julho de 2006, disponível em Ministério da Educação: <http://sinaes.inep.gov.br:8080/sinaes/>
- Losak, J. (1994). *Issues/problems in assessing and documenting institutional effectiveness. research and planning report 94-01. Southern Association of Colleges and Schools*. Acesso em 15 de Agosto de 2006, disponível em Education Resources Information Center: <http://SearchERIC.org/ericdb/ED403812.htm>
- Lowman, J. (2004). *Dominando as técnicas de ensino*. (2 ed.). (H. O. Avritscher, Trad.) São Paulo, SP, Brasil: Atlas.
- Malhotra, N. K. (2001). *Pesquisa de Marketing: uma orientação aplicada*. Porto Alegre, RS, Brasil: Bookman.

- Marsh, H. W. (1984). Students evaluations of university teaching: dimensionality, reliability, validity, potencial biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.
- Ministério da Educação. (2006). *Avaliação Externa de Instituições de Educação Superior - Diretrizes e Instrumento* (Vol. 1). Brasília, DF, Brasil: MEC.
- Ministério da Educação (2006). *Decreto Federal 5.773 - Dispõe sobre o exercício das funções de Regulação, Supervisão e Avaliação de Instituições de Educação Superior e Cursos Superiores de Graduação e Seqüenciais no Sistema Federal de Ensino*. Acesso em 14 de Outubro de 2006, disponível em Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira: http://www.inep.gov.br/download/superior/institucional/Decreto_5773_06.pdf
- Ministério da Educação (2007). *Sistema de Estatísticas Educacionais*. Publicação de 19 de Junho de 2007. Acesso em 11 de Agosto de 2007, disponível em Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira: <http://www.edudatabrasil.inep.gov.br/>.
- Overall, J. V., & Marsh, H. W. (1980). Student's evaluations of instruction: a longitudinal study os their stability. *Journal of Educational Psychology*, (72), 321-325.
- Owlia, M. S., & Aspinwall, E. M. (1988). A framework for measuring quality in engineering education. *Total Quality Management*, 9(6), 501-518.
- Pasquali, L. (1984). Questionário de avaliação de docência (Qad). *Educação e Seleção*, 9, 71-98.
- Pasquali, L. (2001). *Técnicas de Exame Psicológico (TEP) - Fundamentos das técnicas de exame psicológico* (Vol. 1). (A. E. Amaral, Ed.) Brasília, DF, Brasil: Casa do Psicólogo Livraria e Editora.

- Pasquali, L. (2003). *Psicometria: Teoria dos testes na Psicologia e na Educação*. Petropolis, RJ, Brasil: Editora Vozes.
- Pereira, J. C. (1999). *Análise de Dados Qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais* (2 ed.). (S. Biral, Ed.) São Paulo, SP, Brasil: Editora da Universidade de São Paulo.
- Primi, R., & Moraes, I. F. (2005). Validade de um instrumento de avaliação de desempenho de docentes do ensino superior. *Revista Avaliação*, 13(3), 41-49.
- Ramos, J. M. (1986). Validacion de construto en el ambito pedagogico. *Revista Española de Pedagogía*, 174, 535-554.
- Shah, A., Lo, T., Wong, P., & Bradbeer, R. (16-21 de Outubro de 2004). *Student evaluation of teachers in an engineering department*. Acesso em 9 de Novembro de 2005, disponível em Southeastern University College for Engineering Education: http://succeednow.org/icee/Presentations/Pres_340_468.pdf
- Such, J. G. (1999). Análises de tipologías de calidad docente a partir de um cuestionario de evaluación del professorado universitario. *Bordon*, 51(1), 95-113.
- Tejedor, F. J., & Monteiro, M. L. (1990). Indicadores de la calidad docente para la evaluacion del professor universitário. *Revista Española de Pedagogia*, 186, 259-279.
- Tejedor, F. J., Seijas, E. J., & Rodríguez, C. M. (1988). Evaluacion del profesorado universitario por los alumnos en la universida de Santiago. *Studia Pedagogica*, 20(5), 73-134.
- Yaremko, R. K., Harari, H., Harrison, R. C., & Lynn, E. (1986). *Handbook of research and quantitative methods in psychology*. Hillsdale, NJ: Lawrence Erlbaum.

ÍNDICE REMISSIVO POR AUTOR

A	Centra e Gaubatz (1999).....5, 58, 59
Abbot e Perkins (1978), Reardon e Waters (1979) citado por Lowman (2004)6	Ch
Almeida (2002).....2	Churchill (1979), 27
<i>American Educational Research Association, American Psychological Association e National Council on Measurement in Education</i> (1999)26	C
Anastasi e Urbina (2000).....25, 26	Corcoran e Fisher (1987)..... 16
Armstrong (1983).....5	Cronbach (1996)..... 26
Armstrong (1998).....34	Cunha (1999)..... 6
B	D
Basow e Silberg (1987)30	Dillman (1978 citado por Pasquali, 2003) 47
Blanco e Paz (1988)29, 56	Doyle (1977)..... 19, 21, 24, 42, 62
Borich (1986 citado por Such, 1999)19, 21, 24, 42, 62	Doyle (1977, citado por Tejedor & Monteiro, 1990)..... 19, 21, 24, 42, 62
Borich (1986) 19, 21, 24, 42, 62	Duke e Stiggins (1997 citado por Fuentes-Medina & Sánchez, 1999)..... 2
Braskamp, Brandenburg e Cry (1994).....22	E
Buck (1998).....34	Erdle, Murray e Rushton (1985)..... 31
C	F
Casanueva, Periañez e Rufino (1997).....30, 56	Fuentes-Medina e Sánchez (1999).....2, 3
Cashin (1995)27	G
Cashin (1996)10	Gillmore e Greenwald (1999)..... 35
Castro (1992).....33	Greenwald (1997).....33, 34, 43, 53
	Greenwald e Gillmore (1997).....5, 7

Greenwald e Gillmore (1998).....35

H

Hair e colaboradores (2005) .27, 49, 50, 51, 54, 60

Howell e Symbaluk (2001).....36

I

Instituto Nacional de Estudos e Pesquisas
Educaçionais Anísio Teixeira (2004)13

L

Losak (1994)35

Lowman (2004).....5, 6, 7, 57, 58

M

Malhotra (2001).....27

Marsh (1984).....5, 27

Miller (1998, citado por Almeida, 2002).....2

Ministério da Educação (2005)14, 15

Ministério da Educação (2006)10, 14

Ministério da Educação (2007)10, 11, 12, 13

Moomaw (1977 citado por Tejedor, Seijas &
Rodríguez, 1988)4

O

Overall e Marsh (1980)7, 33, 60

Owlia e Aspinwall (1988)29, 56

P

Pasquali (1984).....29, 56

Pasquali (2001).....1, 26

Pasquali (2003).....47

Pereira (1999)54

Primi e Moraes (2005).....17, 24, 43, 53, 62

R

Ramos (1986)27

Rodríguez Espinar (1988).....23

Rodríguez Espinar (1988, citado por Tejedor &
Monteiro, 1990)23

S

Shah e colaboradores (2002)56

Such (1999)19, 21, 24, 42, 62

T

Tejedor e colaboradores (1988)4, 7, 17, 18, 21,
24, 30, 33, 55

Tejedor e Monteiro (1990)2, 3, 4, 5, 6, 17, 19, 21,
22, 23, 24, 42, 45, 57, 58, 59, 62

W

Wilman (1988).....22, 23

Y

Yaremko, Harari, Harrison e Lynn, (1986)25

Yinger (1986, citado por Tejedor e colaboradores,
1988).....18

ANEXOS

Telas do formulário de avaliação docente



Bem-vindo ao Aluno On Line!
Para entrar, você deve fornecer abaixo seu número de matrícula e sua senha.

Aluno ON-LINE

R.A.:

Senha:

[Cadastre/consulte sua senha?](#)

Figura 12. Tela de autenticação do aluno

AVALIAÇÃO DOCENTE						
[DIDÁTICA] [CONTEÚDO] [O PROCESSO DE AVALIAÇÃO] [RELACIONAMENTO] [SUGESTÕES]						
DIDÁTICA DO PROFESSOR (A clareza das explicações nos assuntos tratados, adequação dos métodos de ensino, o aproveitamento do tempo destinado às aulas, o cumprimento do programa apresentado, diversidade de recursos didáticos):						
DIDÁTICA	Excelente	Bom	Regular	Fraco	Péssimo	Não se aplica
Docente 1 DISCIPLINA 1	☐	☐	☐	☐	☐	☐
Docente 2 DISCIPLINA 2	☐	☐	☐	☐	☐	☐
Docente 3 DISCIPLINA 3	☐	☐	☐	☐	☐	☐

[CONFIRMA]

Figura 13. Tela de avaliação do aluno referente à Didática

AVALIAÇÃO DOCENTE						
[DIDÁTICA] [CONTEÚDO] [AVALIAÇÃO] [RELACIONAMENTO] [SUGESTÕES]						
CONTEÚDO (O conhecimento que o professor tem dos conteúdos desta disciplina):						
CONTEÚDO	Excelente	Bom	Regular	Fraco	Péssimo	Não se aplica
Docente 1 DISCIPLINA 1	☐	☐	☐	☐	☐	☐
Docente 2 DISCIPLINA 2	☐	☐	☐	☐	☐	☐
Docente 2 DISCIPLINA 3	☐	☐	☐	☐	☐	☐

[CONFIRMA]

Figura 14. Tela de avaliação do aluno referente ao Conteúdo

AVALIAÇÃO DOCENTE						
[DIDÁTICA] [CONTEÚDO] [AVALIAÇÃO] [RELACIONAMENTO] [SUGESTÕES]						
O PROCESSO DE AVALIAÇÃO (A adequação das avaliações em relação aos conteúdos trabalhados, à objetividade dos critérios de avaliação, a análise e encaminhamentos que o professor faz dos resultados das avaliações visando à melhoria da aprendizagem):						
AVALIAÇÃO	Excelente	Bom	Regular	Fraco	Péssimo	Não se aplica
Docente 1 DISCIPLINA 1	☐	☐	☐	☐	☐	☐
Docente 2 DISCIPLINA 2	☐	☐	☐	☐	☐	☐
Docente 3 DISCIPLINA 3	☐	☐	☐	☐	☐	☐

[CONFIRMA]

Figura 15. Tela de avaliação do aluno referente à Avaliação

AVALIAÇÃO DOCENTE						
[DIDÁTICA DO PROFESSOR] [CONTEÚDO] [O PROCESSO DE AVALIAÇÃO] [RELACIONAMENTO] [SUGESTÕES]						
RELACIONAMENTO (Ao incentivo dado pelo professor para você participar das aulas, à atenção do professor para responder as dúvidas, à postura ética do professor frente aos alunos, colegas, à profissão e à Instituição):						
RELACIONAMENTO	Excelente	Bom	Regular	Fraco	Péssimo	Não se aplica
Docente 1 DISCIPLINA 1	☐	☐	☐	☐	☐	☐
Docente 2 DISCIPLINA 2	☐	☐	☐	☐	☐	☐
Docente 3 DISCIPLINA 3	☐	☐	☐	☐	☐	☐

[CONFIRMA]

Figura 16. Tela de avaliação do aluno referente ao Relacionamento