

# Análise da importância em pesquisas de satisfação através da regressão múltipla: estudo do efeito de diferentes pontos de escala

## *Importance analysis in satisfaction research through multiple regression: effect of different rating scale formats*

Fabiano Larentis<sup>a</sup>  
Cíntia Paese Giacomello<sup>b</sup>  
Maria Emília Camargo<sup>c</sup>

**RESUMO:** O uso de pesquisas de satisfação é fundamental para se captar a voz do cliente, em termos do que ela oferece ao mercado e o que é esperado. Um aspecto de destaque dentro das pesquisas de satisfação é poder identificar quais itens dentro de uma dimensão merecem maior atenção, através da análise da importância. Uma das formas é analisá-la de maneira direta, solicitando ao respondente a indicação de um nível de importância para os mesmos itens avaliados para a satisfação. Outra maneira é a importância derivada, a partir da Análise de Regressão Múltipla. No entanto, o uso de determinado número de pontos de escala e a forma como os pontos são rotulados pode interferir nos resultados de uma regressão. Este estudo objetiva identificar qual a configuração de escala (pontos) é mais apropriada para analisar a importância. Desenvolveu-se uma pesquisa de satisfação junto a 177 alunos de graduação, utilizando o mesmo instrumento com quatro diferentes escalas de resposta: 5, 6, 7 e 11 pontos. Os resultados indicam que a escala de 7 pontos se apresenta superior às outras. Ratifica-se todavia a necessidade de avaliação das circunstâncias de uso das escalas e a necessidade de se considerar qual a configuração mais apropriada.

**Palavras-chave:** Pesquisa de satisfação. Pontos de escala. Análise da importância. Análise de regressão múltipla.

**ABSTRACT:** The use of satisfaction surveys is essential to capture the voice of the customer, in terms of what it offers to the market and what is expected. One relevant aspect within satisfaction surveys is to identify which items within a dimension deserve greater attention, through the analysis of the importance. One way is to analyze it directly, asking the respondent to indicate a level of importance for the same items evaluated for satisfaction. Another way is the importance derived from the Multiple Regression Analysis. However, the use of a number of points to scale and how the points are labeled can interfere with the results of a regression. This study aims to identify which configuration of scale (points) is more appropriate to analyze the importance. It was developed a satisfaction survey among 177 undergraduates, using the same instrument with four different response scales: 5, 6, 7 and 11 points. The results indicate that 7-point scale appears superior to the others. However, it confirms the need to review the circumstances of use of the scales and the need to identify what the most appropriate configuration.

**Keywords:** Satisfaction survey. Scale points. Importance analysis. Multiple regression analysis.

**JEL classification:** M - Business Administration and Business Economics; Marketing; Accounting.

<sup>a</sup> Doutor em Administração (UFRGS). <[flarenti@ucs.br](mailto:flarenti@ucs.br)>.

<sup>b</sup> Doutora em Administração (UFRGS). <[cpaese1@ucs.br](mailto:cpaese1@ucs.br)>.

<sup>c</sup> Doutora em Engenharia da Produção (UFSC). <[kamargo@terra.com.br](mailto:kamargo@terra.com.br)>.



## 1 Introdução

O uso de pesquisas de satisfação é fundamental para se captar a voz do cliente a respeito de como a organização pode ser avaliada em termos do que ela oferece com seus produtos, serviços e marcas e do que é esperado pelos clientes. É comum, em relação a isso, o uso de escalas tipo Likert para a avaliação dos itens que compõem as dimensões/fatores de satisfação analisados.

Um aspecto importante dentro das pesquisas de satisfação é poder identificar quais itens dentro de uma dimensão merecem maior atenção, definindo prioridades de ação para as organizações. Isso exige a avaliação, dentro da pesquisa, da importância que os clientes dão a cada um dos aspectos avaliados, partindo-se do pressuposto que cada item possui um peso particular. Uma das formas, segundo Oliver (2010), é analisar a satisfação e a importância de cada item de maneira pontuada (janela do cliente), porém esta forma de análise gera duplicação do instrumento. Outra maneira, mais apropriada segundo o autor, é analisar a importância de maneira derivada, ou indireta, a partir da Análise de Regressão Múltipla, cujos coeficientes, ou betas, indicariam qual item merece maior atenção, em relação aos outros.

De acordo com Weijters, Cabooter e Schillewaert (2010), a escolha de uma escala de medida particular envolve o número de categorias de resposta (pontos) e sua rotulagem. Segundo os autores, o uso de determinado número de pontos de escala e a forma como os pontos são rotulados pode interferir nos resultados de uma regressão, o que inclui a análise de confiabilidade, geralmente através do  $\alpha$  de Cronbach, e da validade, geralmente através do coeficiente de determinação, ou  $R^2$ . O  $R^2$ , que indica a capacidade de explicação dos itens utilizados em relação à satisfação geral, depende de uma construção criteriosa da escala (definição das variáveis), que considere tanto aspectos considerados importantes pelos executivos, pela revisão da literatura quanto pelos clientes. No entanto, a quantidade de pontos é definida pelo pesquisador, muitas vezes pelo que é comumente utilizado em outras pesquisas, sem saber exatamente qual a consequência dessa escolha.

A Análise de Regressão, ademais, pressupõe coeficientes (betas) que são padronizados,

podendo assumir valores positivos ou negativos (relação inversa). No caso de pesquisas de satisfação, a presença de itens com coeficientes negativos, desde que não reversos, indica um contrassenso, uma vez que, matematicamente, se houver melhora de um item com beta negativo, gerará redução da satisfação geral da dimensão/fator. Isso pode ser decorrência do número de variáveis utilizadas no modelo, da multicolinearidade existente entre elas (HAIR et al., 2005), assim como do número de pontos da escala, que podem permitir ou não maior variabilidade nas respostas.

Desta maneira, este estudo tem por objetivo identificar qual a configuração de escala (pontos de escala) mais apropriada para analisar a importância dentro de pesquisas de satisfação, que permite melhor ajuste aos modelos de regressão. Para tanto, efetuou-se um estudo junto a estudantes de graduação de uma instituição de ensino superior, com a aplicação aleatorizada de quatro modelos de questionários de satisfação, com escalas de resposta variando de totalmente insatisfeito a totalmente satisfeito: 5 pontos, 6 pontos, 7 pontos e 11 pontos (0 a 10). Se avaliaram também atitudes dos estudantes em relação à universidade, com as âncoras de discordo totalmente a concordo totalmente. Este estudo, portanto, se justifica pela disseminação dos estudos de satisfação junto às organizações, incluindo o uso da importância relativa, via análise de regressão múltipla, a interferência no uso de determinada configuração de escala (número de pontos) nos resultados e a possibilidade de surgirem betas negativos, o que dificulta a interpretação dos coeficientes.

O artigo está estruturado da seguinte maneira: após esta introdução, segue o referencial teórico, abordando os temas escalas em marketing e quantidade de pontos de escala, pesquisa de satisfação e escalas e avaliação da importância em estudos de satisfação. Adiante, é apresentado o método e os resultados da pesquisa. Por fim, as considerações finais.

## 2 Referencial teórico

### 2.1 Escalas em marketing e quantidade de pontos de escala

A construção de escalas em Marketing faz parte da rotina de pesquisadores da área. É a partir das escalas que se permite mensurar aspectos associados a comportamentos e

tendências. Quanto a isso, conforme Devlin, Dong e Brown (2003), uma boa escala deve passar pelo teste da validade e da confiabilidade e a sua construção, envolvendo itens, número de categorias e palavras âncora (rótulos), deve permitir facilidade na administração, gerar o mínimo de vieses de resposta e produzir resultados gerencialmente relevantes.

A elaboração de uma escala, além da identificação e validação dos itens/indicadores que farão parte dos construtos envolvidos, envolve decisões quanto ao uso de determinada categoria, ou pontos, de resposta. Conforme Weijters, Cabooter e Schillewaert (2010), a escolha de uma escala de medida particular envolve o número de categorias de resposta (ou pontos de uma escala) e respectivos rótulos. Nessa escolha, define-se o número de pontos, a existência de um ponto médio (o que definirá uma escala ímpar ou par), o uso de escala forçada (sem oferecer opção aos que não sabem responder determinado item) e a rotulagem dos pontos (se todos os pontos ou se somente seus extremos) (MALHOTRA, 2006).

Conforme análise de Weijters, Cabooter e Schillewaert (2010), utilizar somente rótulos extremos facilita a construção de uma escala e esse formato intuitivamente está mais em linha com a presunção de intervalo. Por outro lado, formatos com todos os pontos rotulados facilitam a interpretação tanto para os pesquisadores quanto para os respondentes.

Pesquisadores de marketing, segundo Weijters, Cabooter e Schillewaert (2010), comumente utilizam escalas de cinco, seis ou sete pontos, incluindo os totalmente rotulados (tipo Likert) ou rotulados nos seus extremos. As escalas tipo Likert têm como origem a escala de cinco pontos rotulados desenvolvida por Rensis Likert, que varia de 1 Discordo Totalmente a 5 Concordo Totalmente (MALHOTRA, 2006), e é amplamente utilizada nas ciências sociais principalmente pela facilidade de entendimento por parte dos respondentes.

Por outro lado, tendo em vista uma perspectiva mais gerencial, Wittink e Bayer (2003) indicam que, quanto maior a escala, melhor, pois permite mais opções aos respondentes e maiores oportunidades de melhoria à empresa. Paralelamente, geralmente remete-se ao “mágico” número sete, mais ou menos dois pontos, analisado por Miller (1956). Em análise de Vieira (2011), de mais de cinquenta escalas

em Marketing, em grande parte oriundas de trabalhos nacionais, quase 60% delas apresentou cinco pontos e quase um quarto sete pontos.

Neste aspecto, conforme Leung (2011), a escala Likert é bastante popular, mas a definição de pontos da escala é ainda controversa. Em estudo de diferenças entre quatro, cinco, seis e onze pontos de escala tipo Likert, considerando suas propriedades psicométricas, o autor identificou que ter mais pontos parece reduzir a assimetria, e a escala de 11 pontos, de 0 a 10, tem a menor curtose e é mais próxima à curva normal. Ademais, somente escalas de 6 e 11 pontos apresentaram distribuições normais. Leung sugere o uso da escala de 11 pontos.

Um ponto que deve ser levado em conta na construção de escala é o estilo ou viés de resposta (BAUMGARTNER; STEENKAMP, 2001; WEIJTERS; CABOOTER; SCHILLEWAERT, 2010). Conforme Baumgartner e Steenkamp (2001), estilos de resposta são tendências para responder sistematicamente ao questionário de maneira diferente ao que os itens foram projetados para medir. Podem contaminar respostas observadas (inflacionando ou deflacionando), afetar conclusões a respeito da relação entre escalas (correlações entre os escores dos respondentes) e levar a conclusões seriamente influenciadas. Os autores identificaram sete estilos de resposta, associados basicamente à (não) aquiescência (tendência a concordar ou a discordar), resposta nos extremos e em suas proximidades, resposta no meio e resposta não contingente (relacionada a erro de medida aleatório). Em relação a isso, Weijters, Cabooter e Schillewaert (2010) consideram em seu estudo os estilos de (não) aquiescência, resposta extrema e resposta no meio.

Na relação entre estilos de resposta e número de pontos de escala e sua rotulagem, Weijters, Cabooter e Schillewaert (2010) comprovaram o seguinte: o rótulo de todas as categorias de resposta tende a aumentar aquiescência, uma vez que a clareza e saliência provavelmente reforça a pressão a concordar; rotular leva a menores níveis de respostas extremas, pois tornam os pontos intermediários mais salientes, e de respostas no meio; respostas extremas e de meio se reduzem com a adição de um ponto médio, quando rotulado, mas aumentam a tendência a concordar; resposta extrema se reduz quando mais graduações de pontos são oferecidos; em formatos com os pontos extremos

somente, adicionar mais graduações tende a um aumento de responder no meio.

Ainda, Weijters, Cabooter e Schillewaert (2010) indicam sugestões de uso de escalas para dois grupos de respondentes, levando em conta estilos de resposta. No desenvolvimento de escalas e nas medidas de opiniões, para populações que tendem a apresentar maior habilidade cognitiva, habilidades verbais e experiência com questionário, como estudantes, pode-se reduzir os efeitos dos estilos de resposta e é menos problemático usar escalas com mais pontos de resposta, rotulados, especialmente sete pontos. Para a população em geral, é melhor utilizar menos categorias, especialmente cinco pontos, rotulados.

Para a estimação de relações lineares, tais como correlações, regressões e modelagem de equações estruturais, os autores indicam cinco ou sete pontos com rótulos extremos. Formatos com rótulos extremos, segundo eles, resultam em relações lineares mais fortes entre atitudes e intenções que formatos rotulados. Itens reversos, no entanto, podem se apresentar problemáticos em formatos extremos.

Essas evidências reiteram a relatividade na escolha dos pontos de escala. Mesmo que alguns estudos apresentem resultados favoráveis a escalas com mais pontos, deve-se considerar de forma criteriosa qual é o objetivo da pesquisa e o público-alvo.

## 2.2 Pesquisa de satisfação e escalas

A satisfação é definida como a avaliação pós-consumo relacionada ao atendimento das expectativas em relação ao desempenho de determinado produto ou serviço, resultando em um sentimento de prazer ou desapontamento (KOTLER; KELLER, 2006; BACKWELL; MINIARD; ENGEL, 2008). Isso denota a necessidade de investigar e de compreender como os clientes avaliam empresas e suas marcas, através de pesquisa. Conforme Rossi e Slongo (1998), pesquisa de satisfação de clientes é um sistema de administração de informações que permanentemente capta a voz do cliente, através da avaliação do desempenho da empresa a partir do ponto de vista do cliente, captando a qualidade externa (desempenho) da empresa em seus negócios e indicando caminhos para as decisões futuras de marketing.

Para uma pesquisa que procura mensurar a satisfação dos clientes, Rossi e Slongo (1998)

sugerem que se utilize, além da fase descritiva quantitativa, uma intitulada exploratória qualitativa. Esta fase, dentre outros benefícios, possibilita um instrumento de coleta mais consistente e mais próximo ao fenômeno sendo investigado, uma vez que permite geração de indicadores considerados importantes pelos clientes, pois envolve sua participação. Em relação, a isso, segundo Rossi e Slongo, errar na definição dos indicadores de satisfação é errar na essência da pesquisa: por mais experientes que sejam os executivos, ou por mais consistente que seja a literatura, será necessário dotar a pesquisa de características específicas, devidamente sintonizadas com o perfil do cliente.

Em termos de quantos pontos da escala utilizar, Rossi e Slongo (1998) sugerem cinco pontos, com extremos rotulados (totalmente insatisfeito a totalmente satisfeito). As cinco categorias, segundo eles, possibilitam que o cliente expresse total ou parcial satisfação, estado indefinido entre satisfação e insatisfação e total ou parcial insatisfação.

Por outro lado, Oliver (2010) indica que, como muitos dados relacionados à satisfação não são sujeitos a interpretações estritamente numéricas, escalas de cinco e de sete pontos são mais comuns e de certa maneira padrão de uso, sendo que a de sete é descritivamente mais versátil. Análise de Wittink e Bayer (2003) considera como mais apropriado dez pontos para uma avaliação geral de satisfação, mas dois pontos para os itens específicos.

Oliver (2010), além disso, indica que para propósitos preditivos um mínimo de três pontos de escala é recomendado. Assim que o número de pontos aumenta, excedendo dez, podem surgir problemas, uma vez que consumidores tendem a utilizar subintervalos em escalas muito longas, como restringindo respostas a um intervalo de cinco pontos em uma escala de dez. Estes consumidores estão interpretando, portanto, a escala de uma maneira não conhecida pelo pesquisador. Uma exceção, segundo Oliver, é a escala de onze pontos, de 0 a 10, permitindo uma interpretação percentual, facilmente compreendida pelos respondentes.

De acordo Devlin, Dong e Brown (2003), as escalas de quatro pontos, de expectativas/exigências, e de cinco pontos, de excelência, são as melhores para avaliar satisfação, com  $\alpha$  de Cronbach entre 0,8 e 0,9 e  $R^2$  entre 0,6 e 0,8. Os

autores identificam a escala de quatro pontos melhor que a de cinco pontos.

O estudo de Schmalensee (2003) sobre escalas usa como exemplo de argumentação a de sete pontos. A autora, além disso, discorda da recomendação de Wittink e Bayer, a qual, segundo ela, não é adequada para estudos mais aprofundados de satisfação. Complementando, na obra de Vieira (2011) sobre escalas de Marketing, na seção de satisfação há duas escalas com cinco pontos, duas escalas com sete pontos e uma escala com dez pontos, denotando semelhança com o que é apresentado no livro como um todo.

Tal como colocado na seção anterior, é necessário relativizar o uso de qual escala e de quais pontos. Rossi e Slongo (1998) comentam que não existe uma escala ideal que seja capaz de responder, indistintamente, a todas as pesquisas de medição de satisfação de clientes. É necessário analisar, segundo eles, os propósitos que levam uma empresa a conceber um sistema de avaliação da satisfação de seus clientes e, a partir disso, elaborar o tipo de escala mais apropriado, o que provavelmente se apresente mais confiável e válido.

### **2.2.1 Avaliação da importância em pesquisa de satisfação e o uso de escalas**

Muitas pesquisas de satisfação também avaliam a importância de determinado item. Segundo Oliver (2010), a partir dessa análise, podem ser estabelecidas prioridades de ação, principalmente onde os níveis de importância são superiores aos de satisfação. Uma das formas de identificar importância e satisfação ao mesmo tempo é o instrumento quantitativo, com escalas intervalares, ser duplicado, já que se pede para um mesmo indicador um nível de importância e um nível de satisfação. Com as médias aritméticas cruzadas da importância e da satisfação de um item, encarada como o desempenho do item, tem-se uma matriz de quatro quadrantes, chamada de janela do cliente.

De acordo com Oliver (2010), um problema frequente em pesquisa de importância desse tipo é o nível elevado das médias dos itens, devido aos clientes considerá-los essenciais. Também não é claro se determinada característica é importante com sua presença ou com sua ausência. Problemas recentes e a comunicação de marketing podem tornar mais salientes

determinados itens. Oliver também reforça que indivíduos indicam pontuações de satisfação mais negativas ou positivas a atributos avaliados anteriormente como mais importantes. Por fim, a importância avaliada desta maneira pelo consumidor não acrescenta previsibilidade a modelos de satisfação e indevidamente aumenta o tamanho do instrumento de coleta.

Outra forma é o uso da Análise de Regressão Múltipla, com a existência de indicadores de satisfação geral tratadas como variáveis dependentes, vista como uma maneira de mensuração indireta da importância, procedimento recomendado por Oliver (2010). A Análise de Regressão Múltipla é utilizada para avaliar o impacto estatístico de diferentes variáveis independentes (itens) no comportamento de uma dada variável dependente (satisfação geral com determinada dimensão/aspecto de satisfação). Esta análise revela o quanto uma característica (item) está mais correlacionada com o escore geral quando o efeito de outras variáveis investigadas são controladas (há contribuições distintas, não sobrepostas à satisfação). Essas características são definidas como preditores ou aspectos-chave da satisfação. Assim, ao invés de uma importância pontuada, há uma importância derivada (OLIVER, 2010). Nesse ínterim, em estudo de Baião, Valadares e Claudino (2011), entre a mensuração direta e indireta, não se identificou um método definitivamente superior em termos da análise da importância.

Na Análise de Regressão, deve-se ater principalmente ao  $R^2$ , e aos coeficientes de regressão padronizados, ou betas ( $b$ ). O valor do  $R^2$  representa o percentual da variância da variável dependente explicado a partir do conjunto de variáveis independentes. O  $R^2$  varia de 0 a 1: quanto mais próximo de 1, maior o poder de explicação. O coeficiente de regressão é a medida indireta do nível de importância de uma variável independente (item). Quanto maior o valor do beta, maior o impacto do item isoladamente sobre a satisfação geral. No entanto, Oliver (2010) alerta que a presença de multicolinearidade pode obscurecer a contribuição de uma ou mais variáveis altamente correlacionadas e afetar os resultados.

Para avaliação da validade e confiabilidade dos modelos de regressão, em consideração à importância dos itens sobre a satisfação

geral, Wittink e Bayer (2003) e Devlin, Dong e Brown (2003) indicam a análise do  $R^2$  e do  $\alpha$  de Cronbach, respectivamente. Wittink e Bayer também indicam como possibilidade de avaliar a validade do modelo o menor número de coeficientes de regressão (betas) negativos. Todavia, Gleason, Devlin e Brown (2003) reforçam a necessidade dos coeficientes negativos serem estatisticamente significantes.

Neste caso, coeficientes negativos podem ser decorrência da multicolinearidade, a qual acarretaria em partes maiores de variância compartilhada e níveis mais baixos de variância única da qual os efeitos das variáveis independentes individuais poderiam ser estimados. Assim, a inclusão de uma variável poderia gerar um valor de beta negativo, o que, necessariamente, não implicaria na relação de que à medida que um aumenta o outro diminui (HAIR et al., 2005).

Ademais, Gleason, Devlin e Brown (2003) comentam que um  $R^2$  baixo indica que os itens de um componente (dimensão) inadequadamente cobrem questões importantes do consumidor, e que podem ser devidos à grande quantidade de erro de medida na satisfação geral. Se o  $R^2$  não diminui significativamente com a retirada de variáveis não significantes do modelo, mesmo que negativas, pode-se considerar que seus coeficientes são zero.

Percebe-se, portanto, a adequação do uso da Análise de Regressão para avaliação da importância em pesquisas de satisfação. Por outro lado, identificam-se o uso de  $R^2$  e o menor número de coeficientes negativos significantes como formas adequadas de avaliar a validade de modelos de regressão nesse contexto.

### 3 Método

Este estudo é classificado como descritivo e quantitativo. Para avaliar a adequação da quantidade de pontos de resposta para estimar modelos de regressão em pesquisas de satisfação, foram aplicados quatro instrumentos semelhantes, diferindo somente quanto à quantidade de pontos de escala, a alunos de graduação de cursos de ciências sociais (administração de empresas e ciências contábeis) de uma universidade na região sul do país. Os questionários foram aplicados durante o horário das aulas, em cinco diferentes turmas. Na ocasião foi solicitado que os alunos

responderem as questões e após devolvessem o instrumento ao professor.

#### 3.1 Instruções de coleta dos dados

O instrumento utilizado para coleta de dados foi adaptado de DE TONI et al. (2006). Ele foi composto por 42 questões, divididas em blocos. O primeiro bloco, onde os alunos deveriam avaliar sua satisfação com aspectos do curso de graduação foi composto por 30 questões sobre secretaria, infraestrutura, coordenação do curso, professores e disciplinas. Após cada conjunto de questões foi solicitado para os alunos avaliarem, de forma geral, o aspecto em questão. Este item foi utilizado nas análises de regressão como a variável dependente daquele conjunto de itens.

O segundo bloco, onde os alunos deveriam avaliar a concordância com atributos da Universidade quanto ao curso de graduação foi composto por 12 questões. Este bloco tinha por objetivo estimar os atributos importantes para definição da imagem da Universidade. O último bloco do instrumento solicitou informações pessoais dos alunos que permitissem algum tipo de análise por grupo, como situação no curso, sexo, idade, renda e se é bolsista.

Nos questionários, as mesmas questões foram solicitadas em diferentes quatro escalas de medida, variando a quantidade de pontos de resposta.

- a) Modelo 5 pontos: escala de Likert de cinco pontos, variando de 1 a 5, onde 1 correspondia a Totalmente Insatisfeito e 5 Totalmente Satisfeito. Na escala de concordância, as categorias variavam de 1 a 5, onde 1 correspondia a Discordo Totalmente e 5 a Concordo Totalmente;
- b) Modelo 6 pontos: escala bietápica de seis pontos, onde inicialmente o respondente era solicitado a dizer se estava satisfeito ou insatisfeito com algum item e na seqüência qual seu grau de satisfação: totalmente, muito ou pouco. Na escala de concordância, indicava se discordava ou concordava, e depois o grau;
- c) Modelo 7 pontos: escala do tipo Likert de sete pontos, variando de 1 a 7, onde 1 correspondia a Totalmente Insatisfeito e 7 Totalmente Satisfeito. Na escala de concordância, as categorias variavam de 1 a 7, onde 1 correspondia a Discordo Totalmente e 7 a Concordo Totalmente;

- d) Modelo 11 pontos: escala de onze pontos, variando de 0 a 10, onde 0 correspondia a Totalmente Insatisfeito e 10 Totalmente Satisfeito. Na escala de concordância, as categorias variavam de 0 a 10, onde 0 correspondia a Discordo Totalmente e 10 a Concordo Totalmente.

A escolha de 5, 6 e 7 pontos foi efetuada com base em Oliver (2010) e em Weijters, Cabooter e Schillewaert (2010). A escolha de 11 pontos, por sua vez, em Oliver (2010). Optou-se pelo uso de rótulos em todos os pontos da escala nos modelos de 5, 6 e 7 pontos por envolver opiniões dos entrevistados, tal como sugerido, apesar de se estar lidando também com estimação de modelos lineares, para os quais os autores definem como uma melhor opção rotular apenas os pontos extremos. No caso do modelo de 11 pontos utilizaram-se rótulos nos extremos para facilitar a resposta por parte dos entrevistados.

Os modelos dos questionários foram aleatorizados e distribuídos nas salas de aula, configurando um experimento aleatório (cada uma das salas de aula participantes recebeu os quatro modelos de questionário, distribuídos aleatoriamente). A distribuição aleatorizada elimina a tendenciosidade que poderia resultar de uma distribuição sistemática. O objetivo é distribuir uniformemente os riscos fornecidos pela heterogeneidade das unidades experimentais entre os níveis dos fatores. Uma distribuição aleatorizada simula melhor as condições que são assumidas pelos modelos (WALPOLE et al., 2009). Importante ressaltar que os questionários passaram por um pré-teste com cinco estudantes antes de serem aplicados aos alunos.

### 3.2 Descrição da amostra e coleta de dados

Foram coletados dados relativos a 192 alunos dos cursos de ciências contábeis e administração de empresas. A quantidade de questionários de cada modelo foi 52, 45, 46 e 49, respectivamente modelos de 5, 6, 7 e 11 pontos.

Após a finalização da coleta, passou-se para a limpeza do banco de dados, para garantir que as inconsistências dos dados prejudicassem as análises. Não foram identificados padrões de ausência de resposta nas questões, entretanto questionários que apresentaram mais de 10% de dados faltantes (ROTH, 1994) foram eliminados, resultando em 177 questionários, sendo 47, 39,

44 e 47 questionários dos modelos de 5, 6, 7 e 11 pontos respectivamente.

Foi utilizado o software PASW Statistics 18 para geração das análises estatísticas. Em relação a isso, não foi evidenciada diferença significativa nas quantidades de instrumentos por semestre do aluno (Qui-Quadrado=4,263;  $p=0,641$ ), sexo (Qui-Quadrado=1,693,  $p=0,638$ ), idade (Teste F=1,634,  $p=0,184$ ), renda (Qui-Quadrado=22,029,  $p=0,107$ ) e se o aluno é bolsista Prouni ou não (Qui-Quadrado=5,485,  $p=0,140$ ). Esta análise permitiu afirmar que os modelos de instrumento foram adequadamente distribuídos entre os alunos.

### 3.3 Preparação e análise dos dados

Procedeu-se, na sequência, a análise das não respostas dos questionários que permaneceram na análise. Nesta base, foi observada baixa taxa de não respostas. Hair et al. (2005) afirmam que cabe ao pesquisador identificar padrões nos dados perdidos que possam explicar algum motivo subjacente. A avaliação da localização destas não respostas entre os respondentes sugeriu que não há padrões identificáveis que justifiquem alguma preocupação. Observadas as características das não respostas (baixa frequência e aleatoriedade), optou-se pela substituição dos valores faltantes pelos estimados no procedimento de maximização da expectativa, identificado como EM no aplicativo SPSS. É um método iterativo de dois estágios no qual o estágio E faz as melhores estimativas possíveis dos dados perdidos e o estágio M faz a estimação dos parâmetros, assumindo que os dados perdidos tenham sido substituídos (HAIR et al., 2005).

Para a análise dos *outliers*, ou respostas atípicas, optou-se pelo critério multivariado. Um outlier multivariado possui valores extremos em mais de uma variável ou possui uma configuração de respostas incomum (KLINE, 2005). A distância D2 de Mahalanobis mede a distância de cada caso ao centroide, para uma distribuição com determinada covariância. A distribuição da medida D2 é qui-quadrado com graus de liberdade igual ao número de variáveis. São considerados outliers aqueles casos cujas probabilidades associadas ao valor de D2 forem inferiores a 0,001. Não foram identificados outliers neste conjunto de dados.

Além disso, a realização de análises multivariadas exige que algumas suposições estejam satisfeitas. As principais são normalidade,

linearidade e multicolinearidade. Uma das formas de analisar a normalidade é através dos valores de assimetria e curtose (HAIR et al, 2005). A linearidade é outra suposição importante, pois a maioria das técnicas pressupõe relações lineares entre as variáveis, Pode ser verificada através do estudo dos diagramas de dispersão dos dados. Já a multicolinearidade é observada quando uma variável pode ser explicada pelas outras variáveis da análise (HAIR et al., 2005). Na prática, a multicolinearidade pode ser observada quando as variáveis têm alta correlação entre si, através do estudo da matriz de correlações. Hair et al. (2005) e Kline (2005) sugerem valores aceitáveis até 0,90 e 0,85. As análises de normalidade (univariada), linearidade e multicolinearidade indicam que os dados estão dentro das faixas adequadas e podem ser utilizados nas análises.

Por outro lado, por se tratar de uma coleta envolvendo diferentes modelos de instrumentos, foi necessário padronizar as escalas, convertendo os códigos utilizados para a mesma faixa de variação. Os códigos das categorias utilizadas, inicialmente variando de 1 a 5 (modelo 5 pontos), 1 a 6 (modelo de 6 pontos), 1 a 7 (modelo de 7 pontos) e 0 a 10 (modelo de 11 pontos) foram convertidas para a escala 0 a 10. Assim, no primeiro modelo de questionário o 1, equivalente a "Totalmente Insatisfeito" foi convertido para código zero, o 5, equivalente a "Totalmente Satisfeito" foi convertido para 10 e assim para todos os códigos utilizando interpolação linear. A partir deste ponto, pôde-se comparar os quatro instrumentos pelas análises de regressão.

Após todos os instrumentos estarem na mesma escala, buscou-se identificar se eles eram equivalentes para estimar as médias, independentemente da quantidade de categorias utilizadas para obter as respostas. Assim, foi realizada a análise de variância univariada para todos os 42 itens, utilizando como fator da análise o modelo do questionário. Das 42 questões, quatro apresentaram valor de significância na análise de variância inferior a 0,05: Avaliação do estacionamento ( $p=0,040$ ), Estratégias de ensino utilizadas ( $p=0,043$ ), Imparcialidade nos julgamentos ( $p=0,037$ ) e A Universidade pratica preços adequados àquilo que oferece ( $p=0,036$ ). Pelo fato dos valores de significância serem próximos a 0,05 e as questões serem de diferentes blocos e de caráter independente, aceitou-se que os instrumentos

são equivalentes entre si para estimar médias das questões.

## 4 Resultados

Partindo-se do ponto que os quatro modelos de instrumentos são equivalentes, procedeu-se a análise de regressão linear múltipla para cada aspecto analisado na pesquisa, utilizando-se como variável independente a questão que perguntava a satisfação geral com o aspecto. Este procedimento foi repetido para cada modelo de instrumento.

Em todas as análises foi calculado o coeficiente alfa de Cronbach. Este coeficiente mede a confiabilidade da consistência interna (KLINE, 2005. Hair et al. (2005) afirmam que o limite inferior para o alfa de Cronbach geralmente aceito é de 0,70, apesar de poder diminuir para 0,60 em pesquisas exploratórias – semelhante ao valor indicado por Malhotra (2006).

Procedeu-se, também, a verificação da correlação item-total máxima de 0,8 para cada variável, garantindo que não haja multicolinearidade (KLINE, 2005). Adicionalmente, dentro das estatísticas de colinearidade, se avaliou, para cada variável independente, se os valores de tolerância de cada item eram maiores que o valor resultante de 1 menos o  $R^2$  de cada modelo (LEECH; BARRETT; MORGAN, 2005). Em todas as análises os valores das tolerâncias foram superiores aos de  $(1-R^2)$ , não apresentando problema de multicolinearidade.

Para avaliação da validade dos modelos foi utilizada a estatística do coeficiente ajustado de determinação, ou  $R^2$  ajustado. O coeficiente de determinação mede o ajuste que o modelo de regressão tem para explicar a variação nos dados. Um  $R^2=1$  significa que 100% da variação dos dados foi explicada por aquele modelo de regressão. O  $R^2$  ajustado, por sua vez, considera em seu cálculo o número de variáveis independentes e o tamanho da amostra que cada modelo é baseado. Por este motivo, justifica-se sua utilização na comparação de diferentes modelos (HAIR et al., 2005). Levou-se em conta também o número de coeficientes de regressão negativos e significantes, conforme indicam Gleason, Devlin e Brown (2003).

Os dados foram analisados utilizando-se o método *enter* de entrada dos dados, com o uso da média da variável para os dados

faltantes. Este método foi adotado pelo fato dos pesquisadores desejarem que todas as variáveis permanecessem na análise e que os valores dos coeficientes da regressão fossem estimados para todas variáveis simultaneamente. Ainda, não fazia parte dos objetivos do trabalho identificar um modelo parcimonioso, que mantivesse um bom ajuste, porém com menos variáveis independentes.

A Tabela 1, a seguir, apresenta os principais resultados obtidos com os modelos de regressão adotados. Destacaram-se com asterisco aqueles valores de beta cuja significância apresentou valores inferiores a 0,05.

A análise dos coeficientes Alfa de Cronbach indicou que todos os conjuntos de questões apresentam consistência interna satisfatória. As únicas análises que apresentaram valor do Alfa de Cronbach inferior a 0,8 foram o modelo 5 pontos para a dimensão de qualidade da infraestrutura (0,71) e os modelos 5 pontos e 11 pontos para avaliar a satisfação geral com base na satisfação com cada grupo de questões (0,67 e 0,79, respectivamente). Em todos os demais modelos não houve problemas. Pode-se compreender os valores de Alfa serem mais baixos para esta análise, uma vez que ela foi composta por itens de diferentes aspectos,

Tabela 1 – Modelos de regressão

continua...																
	Likert 5 pontos				Bietápica 6 pontos				Likert 7 pontos				Escala 0 a 10			
	$(\alpha=0,88)$				$(\alpha=0,83)$				$(\alpha=0,91)$				$(\alpha=0,91)$			
<b>Secretaria</b>	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Cordialidade e interesse dos funcionários da secretaria	46	7,34	2,20	0,207	38	7,16	2,06	0,129	41	7,68	1,74	0,195	46	7,52	1,81	0,431*
Capacidade dos funcionários da secretaria em resolver problemas	46	6,90	2,12	0,151	38	7,26	1,64	-0,021	41	7,19	2,38	0,178	43	7,09	2,12	-0,132
Agilidade dos funcionários da secretaria	46	6,79	2,34	0,214	38	7,32	1,76	0,486*	41	7,31	1,83	0,362*	43	7,14	2,08	0,575*
Horário de funcionamento da secretaria	46	7,77	1,69	0,292*	37	7,84	1,44	0,383*	39	7,15	2,38	0,298*	47	7,83	1,83	0,122
<i>Satisfação Geral com a secretaria de graduação</i>	46	7,28	1,96	R <sup>2</sup> 0,49	38	7,68	1,71	R <sup>2</sup> 0,61	41	7,41	1,89	R <sup>2</sup> 0,82	46	7,76	1,86	R <sup>2</sup> 0,81
<b>Qualidade da Infraestrutura disponível</b>	$(\alpha=0,71)$				$(\alpha=0,82)$				$(\alpha=0,88)$				$(\alpha=0,85)$			
	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Salas de aula	47	6,60	2,36	0,173	38	6,89	1,90	0,252*	44	6,47	2,87	0,342*	45	6,84	1,89	0,246*
Estacionamento	47	7,03	2,69	0,298*	36	7,39	2,07	0,047	41	6,37	2,92	0,190	45	7,87	1,95	0,086
Biblioteca	45	8,11	1,86	0,231	38	7,63	2,02	0,144	42	7,48	2,30	0,552*	46	7,98	1,67	0,256*
Laboratórios	41	6,89	2,15	0,205	36	6,61	2,48	0,311*	42	6,13	3,21	-0,074	44	7,05	1,93	0,266*
Espaço de convivência	46	6,85	2,50	0,228	39	6,72	2,13	0,298*	43	6,04	3,12	-0,031	43	6,88	2,15	0,278*
<i>Satisfação Geral com a qualidade da infraestrutura</i>	45	6,89	2,07	R <sup>2</sup> 0,54	37	7,08	1,86	R <sup>2</sup> 0,65	44	6,71	2,51	R <sup>2</sup> 0,67	45	7,27	1,59	R <sup>2</sup> 0,78
<b>Coordenação do Curso</b>	$(\alpha=0,90)$				$(\alpha=0,95)$				$(\alpha=0,97)$				$(\alpha=0,97)$			
	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Cordialidade e interesse da coordenação de curso	47	7,29	2,26	0,357*	39	7,28	2,36	-0,111	43	6,92	2,71	-0,117	47	6,96	2,49	0,217*
Capacidade da coordenação de curso em resolver problemas	45	6,83	2,16	-0,030	38	6,89	2,36	0,305*	44	6,68	2,67	0,064	45	7,11	2,33	0,157
Iniciativa do coordenador de curso	46	6,74	1,96	0,179	39	6,67	2,53	0,477*	43	6,53	2,72	0,171	47	7,02	2,42	0,234*
Agilidade da coordenação de Curso	45	6,83	2,35	0,343	38	6,84	2,26	0,309	44	6,68	2,62	0,755*	46	6,91	2,36	0,402*
<i>Satisfação Geral com a coordenação do curso</i>	46	7,17	2,08	R <sup>2</sup> 0,56	37	6,70	2,50	R <sup>2</sup> 0,85	43	6,43	2,95	R <sup>2</sup> 0,74	47	7,13	2,35	R <sup>2</sup> 0,94
<b>A Universidade X...</b>	$(\alpha=0,86)$				$(\alpha=0,94)$				$(\alpha=0,93)$				$(\alpha=0,94)$			
	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
... trata os alunos com respeito	47	7,82	1,53	-0,077	39	7,64	2,05	0,162	44	7,89	1,98	0,017	47	7,96	1,74	0,004
... oferece ensino de excelência	47	7,07	1,67	0,185	39	7,23	1,69	-0,023	44	7,34	1,78	0,127	47	7,40	1,38	0,133
... pratica preços adequados àquilo que oferece	47	4,36	2,73	-0,002	38	4,79	2,65	-0,036	43	4,47	3,18	-0,164	46	5,85	2,12	0,001
... supre perfeitamente o mercado da região com uma ampla variedade de cursos	47	6,49	2,31	0,318	39	6,51	2,50	0,049	44	6,81	2,52	0,087	47	6,96	1,83	-0,076
... é inovadora	47	6,06	2,32	0,130	39	6,26	2,21	-0,223	44	6,70	2,40	0,219	47	6,89	1,72	0,229
... é uma instituição acolhedora (onde as pessoas se sentem bem-vindas)	47	7,29	1,72	0,034	39	7,28	2,27	0,233	44	7,42	2,03	0,121	47	7,60	1,70	-0,002
... é uma instituição de ensino bem estruturada, bem organizada	47	6,70	1,66	0,077	39	7,38	2,25	0,238	44	7,07	1,99	-0,149	47	7,55	1,54	0,153
... é uma instituição de ensino de tradição	46	7,39	1,17	-0,153	39	7,85	1,74	-0,049	44	7,61	1,86	0,242	47	7,89	1,32	0,193
... é uma instituição que promove o desenvolvimento da região	46	7,39	1,39	0,159	39	7,38	1,84	0,226	44	7,14	2,10	0,056	46	7,93	1,32	0,166
... é a melhor instituição de ensino da região	46	6,47	2,45	-0,135	38	6,84	2,49	-0,019	40	7,27	2,24	-0,016	47	7,21	1,98	0,182
... é uma instituição cujos cursos têm ampla aceitação no mercado	47	7,61	1,80	0,400*	38	7,79	1,79	0,395*	44	7,28	2,28	0,505*	47	7,96	1,50	0,067
<i>A Universidade X tem uma imagem positiva na sua área de abrangência, considerando todas as minhas respostas anteriores neste bloco</i>	47	7,55	1,84	R <sup>2</sup> 0,45	38	7,79	2,12	R <sup>2</sup> 0,63	44	7,67	1,68	R <sup>2</sup> 0,80	47	7,83	1,45	R <sup>2</sup> 0,73

Tabela 1 – Modelos de regressão (conclusão)

	Likert 5 pontos				Bietápica 6 pontos				Likert 7 pontos				Escala 0 a 10			
	$(\alpha=0,92)$				$(\alpha=0,92)$				$(\alpha=0,94)$				$(\alpha=0,92)$			
<b>Professores do curso como um todo</b>	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Domínio do conteúdo e relação com áreas afins	45	6,83	2,16	0,377*	39	7,64	1,99	-0,214	43	7,63	2,26	0,388*	46	7,87	1,38	0,153
Planejamento das aulas	47	7,13	1,95	-0,090	39	7,28	1,92	0,053	44	7,47	2,11	0,350	47	7,83	1,26	-0,042
Clareza e objetividade	47	6,70	2,34	-0,100	39	6,97	2,05	0,330*	43	7,48	2,07	-0,003	47	7,47	1,56	0,361
Relacionamento com os alunos	43	7,33	1,92	0,073	39	7,44	1,71	0,370*	43	7,64	2,19	0,025	47	7,85	1,27	-0,196
Estratégias de ensino utilizadas para promover a aprendizagem	47	6,12	2,26	0,182	39	7,03	2,19	-0,011	44	7,14	2,34	0,299*	47	7,26	1,70	-0,288
Vinculação do conteúdo com a prática	47	6,38	2,38	0,292	39	6,72	2,27	0,173	44	6,95	2,49	-0,244	47	6,79	1,74	0,409*
Imparcialidade nos julgamentos (trata todos da mesma	47	6,60	2,47	0,234	38	7,32	2,14	0,243*	44	7,33	2,22	0,032	47	7,89	1,67	0,479*
Uso do tempo de aula de forma proveitosa	47	6,70	2,56	-0,026	39	7,28	2,08	0,235*	44	7,07	2,32	0,152	47	7,57	1,53	0,189
<i>Satisfação Geral com os professores do curso</i>	47	6,54	2,58	R <sup>2</sup> 0,61	39	7,23	1,93	R <sup>2</sup> 0,84	44	7,55	2,25	R <sup>2</sup> 0,86	47	7,28	1,61	R <sup>2</sup> 0,65
<b>Disciplinas do curso</b>	$(\alpha=0,89)$				$(\alpha=0,87)$				$(\alpha=0,93)$				$(\alpha=0,81)$			
	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Adequação à temática do curso	47	6,81	2,13	0,291*	38	6,95	2,01	0,196*	44	6,98	2,10	0,606*	47	7,49	1,57	0,356*
Conteúdos definidos (ementas)	47	6,81	2,00	0,216	39	7,38	1,95	0,217	44	6,69	2,57	0,205*	47	7,64	1,39	0,246
Importância para você como profissional	47	7,02	2,13	0,493*	39	7,59	2,16	0,577*	44	7,17	2,34	0,207*	47	7,96	1,43	0,044
<i>Satisfação Geral com as disciplinas do curso</i>	47	6,81	2,07	R <sup>2</sup> 0,81	39	7,33	1,91	R <sup>2</sup> 0,80	44	7,16	2,17	R <sup>2</sup> 0,92	47	7,53	1,47	R <sup>2</sup> 0,27
<b>Todos itens de satisfação geral</b>	$(\alpha=0,67)$				$(\alpha=0,87)$				$(\alpha=0,83)$				$(\alpha=0,79)$			
	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$	n	Média	DP	$\beta$
Satisfação Geral com a secretaria de graduação	46	7,28	1,96	0,072	38	7,68	1,71	0,162	41	7,41	1,89	0,267	46	7,76	1,86	0,024
Satisfação Geral com a qualidade da infraestrutura	45	6,89	2,07	0,105	37	7,08	1,86	0,122	44	6,71	2,51	0,007	45	7,27	1,59	0,304*
Satisfação Geral com a coordenação do curso	46	7,17	2,08	0,165	37	6,70	2,50	0,045	43	6,43	2,95	0,170	47	7,13	2,35	0,149
Satisfação Geral com os professores do curso	47	6,54	2,58	0,507*	39	7,23	1,93	0,477*	44	7,55	2,25	0,352*	47	7,28	1,61	-0,041
Satisfação Geral com as disciplinas do curso	47	6,81	2,07	0,213	39	7,33	1,91	0,281*	44	7,16	2,17	0,203	47	7,53	1,47	0,578*
<i>Satisfação geral com o curso de graduação, considerando todos os aspectos avaliados anteriormente</i>	47	6,97	1,64	R <sup>2</sup> 0,58	39	7,13	1,88	R <sup>2</sup> 0,79	44	6,99	1,98	R <sup>2</sup> 0,65	47	7,51	1,25	R <sup>2</sup> 0,63

Fonte: Resultados da pesquisa.

como secretaria, infraestrutura, coordenação, professores e disciplinas. De forma geral, os modelos de escala do tipo Likert de 7 pontos e escala de 0 a 10 (11 pontos) foram os que mostraram melhores resultados.

Quanto aos coeficientes de determinação (R<sup>2</sup> ajustado), percebe-se que nas questões de satisfação com relação à Secretaria, Professores, Disciplinas e Imagem a escala que apresentou melhor ajuste nos modelos de regressão foi a de 7 pontos (Tabela 2). Já nas avaliações de Infraestrutura e de Satisfação Geral, a escala de 7 pontos ficou na segunda posição. Os valores de R<sup>2</sup> ajustado, de forma geral, foram maiores com a escala de 7 pontos.

Considerando que o coeficiente de determinação é uma medida utilizada para saber o quão bom é o modelo de regressão utilizado, ou seja, o quão bem as variáveis independentes explicam a variável dependente, pode-se inferir que esta forma de abordagem mostrou-se ser mais forte para captar a variabilidade geral dos dados. O modelo de escala de avaliação variando de 0 a 10 (11 pontos) foi o segundo melhor modelo, embora tenha apresentado resultados mais baixos de R<sup>2</sup> ajustado no bloco Disciplinas.

Tabela 2 – Ordenação dos modelos em função do valor de R<sup>2</sup> ajustado

	Maior valor (Melhor modelo)		Menor valor (Pior modelo)	
Secretaria	7 pontos	11 pontos	6 pontos	5 pontos
Infraestrutura	11 pontos	7 pontos	6 pontos	5 pontos
Coordenação	11 pontos	6 pontos	7 pontos	5 pontos
Professores	7 pontos	6 pontos	11 pontos	5 pontos
Disciplinas	7 pontos	5 pontos	6 pontos	11 pontos
Satisfação geral	6 pontos	7 pontos	11 pontos	5 pontos
Imagem	7 pontos	11 pontos	6 pontos	5 pontos

Fonte: Resultados da pesquisa.

Outra análise cabível é quanto aos valores estimados dos betas ( $\beta$ ). A quantidade de valores negativos foi semelhante entre os modelos. Além disso, se forem considerados somente as variáveis com betas significantes, ao nível de significância 0,05, percebe-se que não houve valores negativos. Pode-se inferir, também, que a presença de coeficientes de regressão negativos neste estudo não foi devida à multicolinearidade dos dados.

Ademais, observando os betas significantes vê-se que não há padrão de permanência de

uma ou outra variável no modelo. Entretanto, o modelo de escala bietápica com 6 pontos identificou 16 betas como sendo significantes (entre os 40 estimados nos diferentes modelos). O modelo de escala de 0 a 10 identificou 14 betas significantes, a escala do tipo Likert de 7 pontos identificou 12 e a escala de Likert de 5 pontos identificou 8 betas significantes.

Considerando a relação entre o valor de  $R^2$  ajustado e o número de variáveis significantes, novamente percebe-se a superioridade da escala de 7 pontos. Ela conseguiu apresentar melhor ajuste com um número relativamente menor de variáveis.

Portanto, levando em conta qual tipo de escala se apresenta mais adequado para a avaliação da importância relativa dos itens em relação à satisfação geral, pode-se definir como mais apropriada a escala de 7 pontos. Também se deve mencionar, de forma geral, os resultados pertinentes da escala de 11 pontos e da escala de 6 pontos. Importante comentar que a escala de 6 pontos, bietápica, é recomendada em *surveys* por telefone (ALBAUM, 1997). Os resultados, no entanto, possibilitam a recomendação de não se utilizar a escala de 5 pontos, apesar de seu uso disseminado nas pesquisas de marketing.

Importante destacar que o público da pesquisa foi de estudantes de graduação, com maior capacidade cognitiva e verbal e mais experiências em responder pesquisas que o público em geral, conforme indicado por Weijters, Cabooter e Schillewaert (2010). Segundo os autores, estudantes ou outros públicos com maior poder de cognição e experiências em pesquisas, como clientes organizacionais, apresentam dificuldade menor de interpretação de escalas com mais pontos. Assim, considerando a relatividade no uso das escalas, o uso dos resultados deste estudo deve ser considerado com cautela, tendo em vista o público envolvido e as recomendações de Weijters, Cabooter e Schillewaert (2010).

#### 4 Considerações finais

o uso de determinada configuração de escala pode interferir nos resultados de uma análise de regressão, uma das formas utilizadas para avaliar a importância em pesquisas de satisfação. Este estudo teve por objetivo identificar qual a configuração de escala (pontos de escala) mais apropriada para analisar a importância indireta

dentro de pesquisas de satisfação. Em relação a isso, destaca-se que os testes estatísticos realizados garantiram que os quatro modelos testados são equivalentes quanto à quantidade de questionários, perfil dos respondentes e resultados obtidos nas médias de cada questão.

Verificando-se os resultados, pode-se assumir que o modelo que apresentou melhor ajuste pelo valor do  $R^2$  ajustado foi a escala do tipo Likert de 7 pontos. O segundo melhor modelo foi a escala de satisfação variando de 0 a 10. O modelo de escala bietápica, embora indicado na literatura como uma possibilidade para facilitar o entendimento, não se mostrou superior aos modelos já citados, embora tenha apresentado resultados satisfatórios. Neste trabalho, o modelo que apresentou os piores ajustes foi o modelo de Likert de 5 pontos.

Quanto aos valores dos coeficientes padronizados (betas), todos os quatro modelos mostraram-se semelhantes, não indicando variáveis com betas negativos com significância estatística ao nível 0,05. Porém quando analisadas as quantidades de variáveis cujos betas foram significantes, a escala de 7 pontos consegue manter um melhor ajuste com menos variáveis. Do ponto de vista gerencial, este seria um ponto a favor, pois permitiria focar esforços para melhoria da satisfação dos clientes tendo que atuar em uma quantidade menor de aspectos.

Este estudo, portanto, ratifica a necessidade de se avaliar as circunstâncias de uso das escalas e a necessidade de se considerar qual a configuração mais apropriada (ROSSI; SLONGO, 1998; WEIJTERS; CABOOTER; SCHILLEWAERT, 2010), em função das diferenças presentes entre os modelos testados. Tanto quanto a definição dos itens de um construto, a escolha da configuração de escala mais adequada deve ser tratada de maneira criteriosa.

Como futuros estudos, se sugere o questionamento aos respondentes sobre seu grau de preferência na adoção de uma ou outra escala, apesar de sua complexidade em termos de aplicação (uma opção é se desenvolver um estudo qualitativo). Outra sugestão de futuros estudos é a aplicação de pesquisas de satisfação junto ao público em geral, devido às peculiaridades em capacidades verbal e de cognição e experiências em responder pesquisas, bem como o uso de mais escalas apenas com os rótulos extremos.

## Referências

- ALBAUM, G. The Likert Scale Revisited: An Alternate Version. *Journal of the Market Research Society*, v. 39, n. 2, p. 331-348, Apr. 1997.
- BAIÃO, A. L.; VALADARES, J. L.; CLAUDINO, L. P. A importância dos atributos na qualidade percebida em serviços: um estudo comparativo entre os métodos da importância declarada e derivada. In: ENCONTRO DA ANPAD, 35., 2011, Rio de Janeiro. *Anais...* Rio de Janeiro: ANPAD, 2011.
- BAUMGARTNER, H.; STEENKAMP, J. E. M. Response styles in marketing research: a crossnational investigation. *Journal of Marketing Research*, v. 38, p. 143-156, May 2001.
- BLACKWELL, R. D.; MINIARD, P. W.; ENGEL, J. F. *Comportamento do Consumidor*. São Paulo: Cengage Learning, 2008.
- DE TONI, D. et al. Análise da satisfação com Instituições de educação superior e imagem: comparando Instrumentos. In: ENCONTRO DA ANPAD, 30., 2006, Salvador. *Anais...* Salvador: ANPAD, 2006.
- DEVLIN, S. J.; DONG, H. K.; BROWN, M. Selecting a scale for measuring quality. *Marketing Research*, p. 13-16, fall 2003.
- GLEASON, T. C.; DEVLIN, S. J.; BROWN, M. In search of the optimum scale. *Marketing Research*, p. 25-29, fall 2003.
- HAIR Jr, J. et al. *Análise multivariada de dados*. Porto Alegre: Bookman, 2005.
- KLINE, R. B. *Principles and practice of structural equation modeling*. 2. ed. New York: Guilford Press, 2005.
- LEECH, N. L.; BARRET, K. C.; MORGAN, G. A. *SPSS for intermediate statistics: use and interpretation*. 2. ed. New Jersey: Lawrence Erlbaum Associates, 2005.
- LEUNG, S. O. A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point likert scales. *Journal of Social Service Research*, v. 37, n. 4, p. 412-421, July 2011.
- MALHOTRA, N. K. *Pesquisa de marketing: uma orientação aplicada*. 4. ed. Porto Alegre: Bookman, 2006.
- MILLER, G. A. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, v. 63, p. 81-97, 1956.
- OLIVER, R. L. *Satisfaction: a behavioral perspective on the consumer*. 2. ed. New York: M. E. Sharpe, 2010.
- ROSSI, C. A. V.; SLONGO, L. A. Pesquisa de Satisfação de Clientes: o Estado-da-Arte e Proposição de um Método Brasileiro. *Revista de Administração Contemporânea (RAC)*, v. 2, n. 1, p. 101-125, jan.-abr. 1998.
- ROTH, P. L. Missing data: a conceptual review for applied psychologists. *Personnel Psychology*, Durham, v. 47, n. 3, p. 537-560, autumn 1994.
- SCHMALENSEE, D. H. The "Perfect" Scale. *Marketing Research*, p. 23-25, fall 2003.
- VIEIRA, V. A. *Escalas em Marketing*. São Paulo: Atlas, 2011
- WALPOLE, R. E. et al. *Probabilidade e estatística para engenharia e ciências*. 8. ed. São Paulo: Cengage Learning, 2009.
- WEIJTERS, B.; CABOOTER, E.; SCHILLEWAERT, N. The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, v. 27, n. 3, p. 236-247, 2010.
- WITTINK, D. R.; BAYER, L. R. The Measurement Imperative. *Marketing Research*, p. 19-23, fall 2003.

Recebido: 30/01/2013

Aceito: 19/11/2014

**Autor correspondente:**

Fabiano Larentis  
 Universidade de Caxias do Sul – UCS  
 Rua Francisco Getúlio Vargas, 1130, Bloco F, 4º andar  
 CEP 95070-560 Caxias do Sul, RS, Brasil  
 E-mail: <[flarenti@ucs.br](mailto:flarenti@ucs.br)>