
Desenvolvimento de um instrumento para mensurar a satisfação do usuário de sistemas de informações

José Dutra de Oliveira Neto
Edson Luiz Riccio

RESUMO

Neste trabalho, objetiva-se desenvolver um instrumento para mensurar, empiricamente, a satisfação do usuário de Sistemas de Informações. Para tanto, foi desenvolvido um instrumento para medir a satisfação do usuário em relação a aplicativos específicos elegíveis pelo próprio pesquisado pelo método *survey*. Uma amostra de 143 respostas de participantes do MBA da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto (FEARP) da Universidade de São Paulo (USP) foi utilizada e o resultado aponta um instrumento (de 5 pontos na escala tipo Likert) cujo coeficiente de consistência interna medido por meio do coeficiente alfa de Cronbach foi de 0.84 e cuja validação empírica foi feita pela análise fatorial e depois confirmada pela análise multidimensional em um conjunto variado de aplicativos. Foram identificados quatro componentes da satisfação do usuário: praticabilidade, disponibilidade, precisão e adequação da informação. Essa pesquisa sugeriu, também, que a satisfação do usuário, quando decomposta em quatro componentes, pode ser base para uma ferramenta de diagnóstico na implementação de sistema de informação, aumentando suas chances de sucesso.

Palavras-chave: satisfação do usuário, sistemas de informações.

1. INTRODUÇÃO

A importância dos sistemas de informações (SIs) nas empresas, aliada a um orçamento crescente na área de Tecnologia de Informação (TI), significa que o seu insucesso, decorrente de falhas ou desuso, poderá gerar grandes perdas para a empresa. Em pesquisa coordenada por Caldas & Wood Jr. (1999) foram constatados diversos problemas na implementação do *Enterprise Resource Planning* (ERP) no Brasil, dos quais se destacam: 45% não obtiveram aumento de produtividade, 43% não diminuíram o ciclo produtivo, 40% não melhoraram os serviços ao cliente, sendo que 36% focaram o lado tecnológico e apenas

Este artigo foi apresentado no *Thirteenth Asian-Pacific Conference on International Accounting Issues* — realizada de 28 a 31 de outubro de 2001 no Rio de Janeiro/RJ, Brasil.

Recebido em 29/novembro/2001
Aprovado em 24/abril/2003

José Dutra de Oliveira Neto é Professor Doutor do Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo (CEP 14040-900 — Ribeirão Preto/SP, Brasil).
E-mail: dutra@usp.br

Endereço:
Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Departamento de Contabilidade e Atuária
Avenida Bandeirantes, 3900
14040-900 — Ribeirão Preto — SP

Edson Luiz Riccio é Professor Livre-Docente do Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (CEP 05508-900 — São Paulo/SP, Brasil).
E-mail: elriccio@usp.br

24% deram atenção ao problema humano na implementação em uma amostra de 85% de empresas transnacionais. Isso pode indicar a necessidade de novos instrumentos que não apenas auxiliem na mensuração do sucesso dos SIs, mas também contribuam para aumentar a probabilidade de sucesso na implementação desses sistemas.

2. PROXIES DO SUCESSO DO SISTEMA DE INFORMAÇÕES

Quanto aos critérios para analisar a questão do sucesso de um SI, aceitou-se a divisão em seis categorias, conforme Delone & Mclean (1992). Essas categorias são: qualidade do sistema, qualidade da informação, uso, satisfação do usuário, impacto individual e impacto organizacional.

Todas essas categorias são amplas e a escolha de uma delas ainda permite várias opções. Por exemplo, a categoria **satisfação do usuário** permite tanto a análise de facilidade de uso (*user friendly*), tão cara aos especialistas de computação, como a amplitude de confiança dos usuários quanto ao atendimento das suas necessidades de informação (IVES & OLSON, 1984).

Poder-se-ia argumentar que a medida econômica ideal de qualquer sistema corporativo é seu impacto em termos monetários, mas essa medida foi considerada impraticável por Galleta & Lederer (1989).

As medidas de sucesso do SI tendem a ser *proxies* do desempenho da empresa, e a escolha da *proxy* corresponde à primeira decisão importante de qualquer estudo como o aqui proposto. Assim, optou-se por utilizar a satisfação do usuário, tendo como critério específico a atitude do usuário em relação ao sistema aplicativo.

Essa decisão segue uma tradição da literatura de utilizar **satisfação do usuário** como uma medida de sucesso (IGBARIA & NACHMAN, 1990; IVES, OLSON & BAROUDI, 1983; MELONE, 1990; HWANG & THORN, 1999; GELDERMAN, 1995 *apud* GELDERMAN, 1998).

Uma das razões para tal popularidade são as crenças de que o objetivo do SI é ajudar o usuário a ter melhor desempenho e de que o usuário satisfeito tem desempenho superior ao dos insatisfeitos. Essas crenças, inicialmente expostas por Bailey & Pearson (1983), são por eles compartilhadas e parte importante do processo que levou à decisão de se concentrar na **satisfação do usuário**. Aliás, essa categoria, assim como **uso**, são as mais utilizadas na literatura como *proxies* das medidas relacionadas a desempenho (AMOROSO & CHENEY, 1991; EIN-DOR & SEGEV, 1982; IGBARIA & NACHMAN, 1990; SCHIFFMAN, MEILE & IGBARIA, 1992).

Neste trabalho, não se considerou a categoria **uso**, pois em situações de **uso mandatório** obviamente esse critério tem valor limitado, como, aliás, já foi apontado por Melone (1990) e Udo (1992).

De acordo com a literatura consultada, o assunto tem sido muito pesquisado nos últimos anos com relativo sucesso. A

seguir, descrevem-se as posições e críticas de alguns autores da área.

Ainda de acordo com Melone (1990), a falta de concordância na definição conceitual do construto referente à satisfação do usuário tem levado a uma situação em que existem muitas operacionalizações e igual número de definições, sendo a maior parte sem fundamento teórico. Tampouco existe consenso quanto à forma de mensuração, conforme constatado nos instrumentos de Ives, Olson & Baroudi (1983) e Doll & Torkzadeh (1988), entre outros. Como consequência, as escalas utilizadas são construídas sem embasamento teórico, na opinião de Woodroof & Kasper (1998).

Segundo Gelderman (1998), a validação dos achados da pesquisa em que essas medidas são usadas para operacionalizar o sucesso do SI depende da validade dessas medidas. Conseqüentemente, o desenvolvimento de fundamentos teóricos e empíricos para sua aplicação é de fundamental importância.

Straub (1989) e Kim (1990) afirmam que, com os instrumentos validados, os pesquisadores podem medir o mesmo construto de pesquisa da mesma maneira, além de permitir maior clareza na formalização e na interpretação das questões de pesquisa.

Como se pode observar na literatura, a validação é fundamental na construção de instrumentos de mensuração de atitude, e sua ausência pode implicar comprometimento das interpretações. Para que se possa desenvolver um instrumento fundamentado teoricamente, e validado, propõe-se a utilização de uma metodologia apropriada, conforme será descrita a seguir.

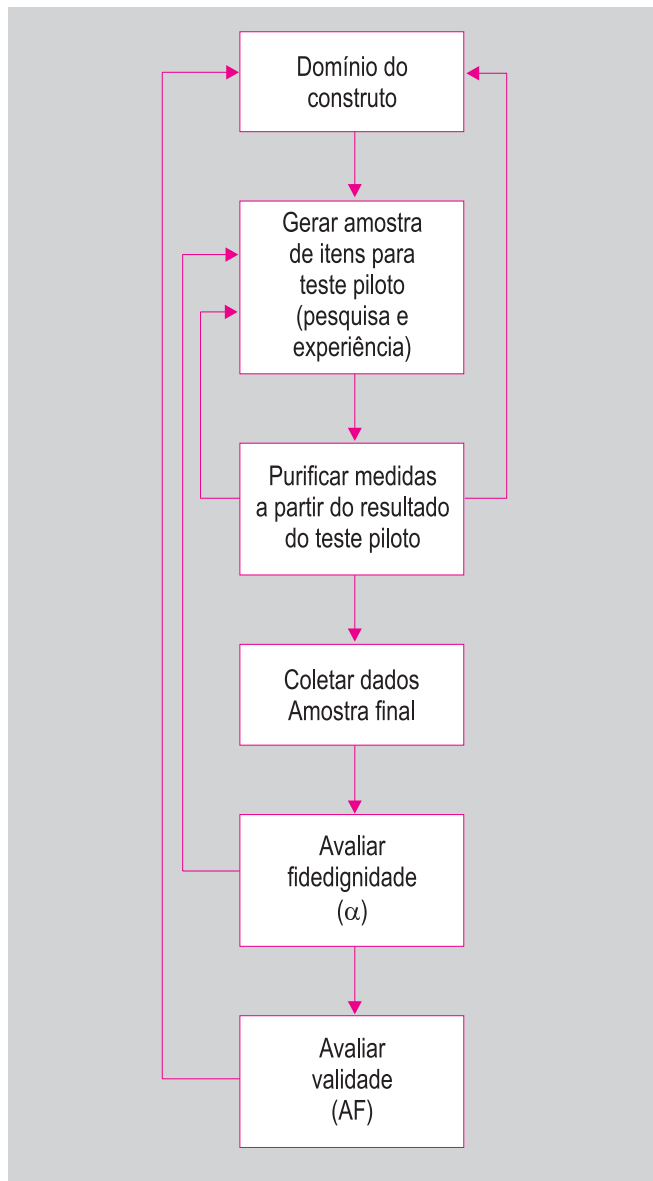
3. METODOLOGIA DE PESQUISA

A metodologia adotada neste trabalho para o desenvolvimento de um instrumento *off-line* com o propósito de medir a satisfação do usuário foi decorrente do modelo de Churchill (*apud* JOSHI, 1989). O *software* utilizado nessa metodologia foi o SPSS® v8. Na figura da página seguinte pode ser observada a metodologia utilizada.

3.1. Etapa 1 — Domínio do construto

Os objetos deste trabalho são os aplicativos específicos — *software* de aplicação específica — utilizados em determinadas atividades e que têm como objetivo lhes dar suporte.

A fim de proporcionar a generalização do trabalho, faz-se necessário envolver os usuários de aplicativos específicos e identificar uma amostra de sujeitos diferenciados representantes de grande número de empresas, profissões e funções. Para tanto, foram escolhidos os alunos dos cursos de Pós-Graduação *Lato Sensu* da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto (FEARP) da Universidade de São Paulo (USP) como amostra deste trabalho, embora se reconheça o provável risco de um certo viés (*bias*) que deve ser



Etapas da Metodologia Utilizada

Fonte: Extraída e adaptada de Joshi (1989, p.347 — baseada em documento de G.A. Churchill).

levado em conta nas avaliações decorrentes da utilização dessa ferramenta.

Outros dois aspectos foram decisivos na escolha da amostra:

- menor rejeição para a aplicação do questionário elaborado por um professor da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP);
- nível de pós-graduação, em que existem profissionais mais atualizados e com maiores possibilidades de utilização de *software* de aplicativos específicos.

A análise fatorial será utilizada como uma ferramenta de análise teórica, com o objetivo de identificar as dimensões da

satisfação do usuário em relação aos aplicativos específicos. Por meio da análise fatorial poder-se-á:

- observar se as percepções dos usuários podem ser agrupadas;
- reduzir o número de itens ou variáveis, inicialmente muito grande.

3.2. Etapa 2 — Geração da amostra de itens

A satisfação do usuário é sua atitude em relação ao sistema aplicativo. Para isso, é preciso identificar os objetos que afetam tais atitudes e, então, utilizar um instrumento para medidas. Para garantir que grande número de objetos relevantes fosse analisado, diversos autores foram pesquisados, como Doll & Torkzadeh (1988), Bailey & Pearson (1983), Ives, Olson & Baroudi (1983), Etezadi-Amoli & Farhoomand (1991), Igbaria & Nachman (1990), Palvia (1996) e Kim, Suh & Lee (1998).

Além da compilação dos itens presentes nas pesquisas existentes — com predominância no instrumento validado por Doll & Torkzadeh (1988) —, foram feitas entrevistas com alguns alunos do MBA — Administração da FEARP-USP, para efetuar o levantamento das crenças modais acerca de aplicativos, o que ampliou o questionário inicial para 24 itens.

Esses 24 itens passaram pela avaliação de diversos alunos dos cursos de Pós-Graduação *Lato Sensu* da FEARPUSP, assim como de funcionários administrativos dessa mesma faculdade com conhecimentos diferenciados em relação a aplicativos. Essa avaliação visou a verificar a compreensão das pessoas com relação à informação requerida. Os sujeitos foram orientados a identificar itens ambíguos, não-claros ou confusos. Foi solicitado, também, que fizessem uma breve descrição oral de sua interpretação de cada item. As informações coletadas foram gravadas e analisadas posteriormente.

Foram constatadas, neste trabalho, evidências de que as dimensões culturais influenciam o comportamento dos respondentes em um questionário que mede a satisfação do usuário. Segundo Litwin (1995), ao se projetar um novo instrumento ou utilizar um instrumento já existente em amostra de diferentes nacionalidades, deve-se antes efetuar a validação cultural na amostra-alvo e depois aplicar. Mesmo sendo possível efetuar a tradução, pode-se não derivar a dimensão esperada. Um conceito bem difundido em uma cultura pode não existir em outra.

Alguns aspectos importantes foram constatados na construção do questionário. Por exemplo, a palavra **sistema**, presente em diversos itens do questionário, foi substituída pela palavra **aplicativo** em razão da divergência de seu significado para as pessoas entrevistadas. Isso decorre, talvez, da tradução e da não-familiaridade de muitos com os termos de informática. O mesmo efeito foi constatado por Mathieson & Ryan (1994) quanto à palavra **sistema** que foi considerada ambígua, o que poderia influenciar os resultados dos instrumentos de satisfação em função da avaliação individual do significado, denominado **efeito da variação de definição**. Os termos *user friendly*

e *timeliness* também tiveram em sua tradução a maior diversidade de entendimento, o que provocou a reformulação das perguntas do questionário, de modo mais específico para manter o mesmo significado, corroborando o trabalho realizado por Hofstede (1980), no qual foi constatado que pessoas de diferentes culturas lidam com a modernização tecnológica de maneiras diferentes e, ainda, que valores culturais preexistentes podem aumentar essa diferença. Mostrou, também, que indicadores culturais da América do Norte diferem dos da América Latina.

O termo **relevância** tem, no Brasil, uma conotação de importância diferente da assumida nos Estados Unidos, onde significa **pertinência**. Os termos **precisão** e **acurácia** (*precision*, *accuracy*) têm significados diferentes para grande parte dos entrevistados. A linguagem dos itens foi refinada para tornar as afirmações mais claras e menos ambíguas. Não foram colocadas novas definições para os termos no questionário por ser um procedimento ineficaz, dado o complexo processo de formação de atitudes.

O veículo para medir a satisfação do usuário em relação a esses itens foi obtido na literatura de Psicologia. Foi utilizada uma escala tipo Likert, de 5 pontos, variando de “pouca importância” até “extrema importância”. A opção pela escala de 7 pontos não foi adotada, uma vez que em metanálise realizada por Chau (1999) ficou demonstrado que as escalas de 7 pontos não obtiveram resultados mais significativos do que as escalas de 5 pontos. O questionário completo pode ser encontrado no Apêndice.

Findo esse trabalho, o questionário foi completamente remodelado e ficou pronto para o pré-teste. O resumo das características do questionário está descrito no quadro a seguir. O questionário final foi condensado em apenas uma página, mantendo o tamanho adequado para uma boa leitura.

Resumo das Características do Questionário

Características da Escala
Número de itens = 16 Tipo de escala = tipo Likert Número de pontos na escala = 5
Procedimentos de Desenvolvimento da Escala
Origem da escala = emprestada e modificada Geração de itens = revisão da literatura e entrevistas
Características da Amostra
Tamanho da amostra = 175 Tipo de respondente = alunos de MBA Método de coleta de dados = questionário com supervisão

3.3. Etapa 3 — Teste piloto

A escala do instrumento foi pré-testada em uma turma de Pós-Graduação *Lato Sensu* de Contabilidade e Controladoria da Fundação para Pesquisa e Desenvolvimento da Administração, Contabilidade e Economia (Fundace), instituição conveniada com a FEARP. Vinte e quatro sujeitos — o correspondente a 13% da amostra final — com perfis semelhantes ao da amostra final (a amostra final refere-se a outras nove turmas de Pós-Graduação *Lato Sensu* de Contabilidade e Controladoria da Fundace) preencheram o questionário composto por 24 itens. Para essa amostra, obteve-se um valor do alfa de Cronbach de 0.75 no teste de confiabilidade.

Segundo Litwin (1995), um dos estágios mais importantes e críticos no desenvolvimento de um novo instrumento é o teste piloto, ou pré-teste, em que os erros são passíveis de correção. Problemas tipográficos, de compreensão e culturais podem interferir no resultado final da pesquisa.

É importante ressaltar a representatividade dessa amostra, uma vez que terá impacto no instrumento resultante deste trabalho. O resultado do teste piloto do instrumento com uma turma de MBA da Fundace — com um pequeno grupo de sujeitos semelhantes à amostra final — sugere que o instrumento conta com propriedades suficientes para dar continuidade à pesquisa.

3.4. Etapa 4 — Purificação das medidas

Churchill (*apud* DOLL & TORKZADEH, 1991) descreve a necessidade de purificar as medidas antes de realizar a análise fatorial. O mesmo autor afirma que, se a análise fatorial for realizada antes da purificação, existe tendência para maior número de dimensões que são fruto de “itens irrelevantes” e que podem comprometer a interpretação da análise fatorial. Isso é feito para que o instrumento contenha apenas os itens que meçam a satisfação, e não itens que sejam a causa da satisfação (como no instrumento de Bailey & Pearson (1983), em que existe o item **envolvimento** que pode ser causal com referência à satisfação).

O número de questões derivadas das etapas anteriores (24) não atendia à proporção ideal mínima de 10:1, conforme Kerlinger (1980), em função da amostra provável de 170 alunos. Para adequar o instrumento à amostra provável, fez-se necessário adotar e identificar alguns critérios racionais para a condução do processo de ajuste ou purificação dos dados. O número adequado de questões — que no caso seria 16 — já contabiliza uma pequena margem de segurança para eventuais respostas inválidas ($10 \times 16 = 160 < 170$).

Os critérios de eliminação adotados pelos pesquisadores foram:

- número de correlações acima de 0,40 (analisando a inter-relação entre os itens);
- importância do item para o trabalho (percepções dos pesquisadores);

- análise gráfica (analisou-se a distribuição das respostas para cada item).

Como resultado conjunto da utilização desses critérios, foi elaborado um questionário final com 16 questões, concebido em apenas uma página para facilitar a tarefa dos respondentes.

Uma preocupação comum entre os pesquisadores é que o resultado desse processo possa ter como consequência a eliminação de alguns itens relevantes. O teste piloto e o processo de eliminação, segundo Doll & Torkzadeh (1991), são procedimentos-padrão no desenvolvimento de escalas tipo Likert. Nesse processo, pode-se enunciar algumas razões para essas eliminações:

- os itens incluídos devem medir a satisfação e não os fatores que possam causá-la. Utilizou-se aqui o critério “importância do item para o trabalho” da metodologia adotada;
- o modelo utilizado permite um método racional para a eliminação de itens. No modelo proposto todos os itens devem contribuir de maneira uniforme e, portanto, as respostas devem estar inter-relacionadas. Utilizou-se aqui o critério “número de correlações acima de 0,40” da metodologia adotada;
- Fishbein & Ajzen (*apud* DOLL & TORKZADEH, 1991) descrevem a necessidade de eliminar itens na construção de escalas tipo Likert. Os itens devem indicar atitudes favoráveis ou não-favoráveis em relação ao objeto em questão. Se um item é ambíguo ou indica atitude neutra deve ser eliminado. Foi utilizado o critério “análise gráfica” da metodologia adotada.

Para corroborar essa linha de pensamento pode-se destacar a afirmação de Hofstede (1980), o qual assegura que, ao incluírem-se variáveis sem inter-correlações na análise fatorial, podem ser identificados fatores fortes que não interferem nas interpretações dos resultados.

A forma inicial do questionário de 24 questões sofreu algumas alterações em decorrência do teste piloto. As alterações foram:

- oito questões foram eliminadas segundo os critérios acima adotados;
- inverteu-se a ordem em cada uma das questões para uma escala mais natural, isto é, de menor intensidade para maior intensidade no lado direito;
- eliminou-se da escala o item “muito pouca ou nenhuma importância”, por não ter sido registrada ocorrência alguma dele no pré-teste. Para manter os 5 pontos na escala e contribuir para a diluição das respostas, foi incluído um novo item na escala entre “importância moderada” e “pouca importância”: “uma certa importância”;
- o questionário final enquadrado-se em apenas uma página para facilitar a aplicação.

Considerando existirem diversos aplicativos em uso nas empresas, foi solicitado aos respondentes que especificassem

apenas o principal aplicativo utilizado (escolhido pelo próprio sujeito).

Foram considerados apenas os usuários que utilizavam algum tipo de aplicativo específico para desenvolver suas atividades. Os aplicativos genéricos como Excel e Word não foram considerados neste trabalho.

Antes de aplicar o questionário definitivo no campo, três funcionários administrativos da FEARP-USP preencheram o questionário para o teste final do instrumento e para identificar o tempo necessário para o preenchimento. Nenhuma modificação foi necessária e o tempo máximo foi de 15 minutos. Os respondentes foram informados sobre o tempo necessário para o preenchimento.

Finda a etapa de purificação decorrente da aplicação do teste piloto, o questionário com 16 questões foi aplicado na amostra final completa.

3.5. Etapa 5 — Coleta de dados (amostra final)

Inicialmente, foram testadas nove turmas de alunos dos cursos de Pós-Graduação *Lato Sensu* da FEARP-USP. Para aumentar a amostra, foi incluída uma turma do curso *Lato Sensu* de Bebedouro (IMESB) sob a coordenação da FEARP-USP.

Todos os respondentes devolveram os questionários após 15 minutos. Conseguiram-se 143 respostas válidas de um total de 175; 32 foram desconsideradas por não se enquadrarem no perfil exigido pelo trabalho, ou seja, por não utilizarem um aplicativo específico ou por deixarem questões em branco. A proporção de 9:1 é considerada adequada, o que permite obter resultados mais confiáveis.

Os sujeitos que compõem a amostra válida (143) pertencem a um grupo bem diversificado, como pode ser constatado nas tabelas de 1 a 6.

Tabela 1

Cargos Predominantes

Cargo	Porcentagem
Gerente	35
Analista	17
Diretores ou Supervisores	15

Tabela 2

Áreas Predominantes

Área	Porcentagem
Administrativa	24
Contabilidade	19
Financeira	14
Comercial	13

Tabela 3**Graduação — Distribuição Predominante**

Graduação	Porcentagem
Administração	25
Contabilidade	23
Engenharia	15
Economia	12

Tabela 4**Perfil dos Respondentes com Relação a Pós-Graduação**

Pós-Graduação	Porcentagem
Graduação apenas	71
Curso de especialização concluído	18

Tabela 5**Sistema Utilizado pelo Respondente — Distribuição Predominante**

Sistema Utilizado	Porcentagem
ERP	30
Sistemas Contábeis	25
SSD (Sistemas de Suporte à Decisão)	15

Tabela 6**Horas Semanais de Uso do Aplicativo — Valores Predominantes**

Horas que Utiliza o Aplicativo	Porcentagem
De 5 até 10 horas	23
De 10 até 20 horas	21
Acima de 30 horas	18

Segundo Freitas *et al.* (2000), a validade e a confiabilidade são requisitos essenciais para o desenvolvimento de um instrumento. Neste trabalho, mediu-se a confiabilidade por meio do teste de consistência interna (alfa), que é um indicador do grau em que os itens que constam do questionário medem o mesmo assunto. Além de determinar a confiabilidade da escala, deve-se determinar sua validade. A validade indica o grau em que um dado indicador mede o que foi projetado para medir, e não outro fenômeno relacionado. Segundo Nunally (1978, *apud* CARMINES & ZELLER, 1979), a análise fatorial deve ser utili-

zada como uma ferramenta de análise teórica, e pode ser útil para estimar a validade de medidas empíricas. Portanto, neste trabalho, utilizou-se o teste de consistência interna (alfa) para a confiabilidade e a análise fatorial para estimar a validade das medidas.

3.6. Etapa 6 — Testando a confiabilidade

O teste de confiabilidade é imperativo, visto que se deve medir o desempenho de um instrumento em uma dada população e, ao mesmo tempo, evitar o agrupamento de questões aparentemente relevantes (LITWIN, 1995). A confiabilidade obtida pela amostra final está descrita no quadro a seguir.

Confiabilidade Geral (Gerada pelo SPSS)

Coefficiente de Confiabilidade 16 itens
Alfa de Cronbach = 0,8493
Alfa de Cronbach Item Padronizado = 0,8530

A confiabilidade deve melhorar, claramente, com o aumento da amostra. Dentro de certos limites (NUNALLY, 1978, *apud* CARMINES & ZELLER, 1979), a extensão do teste deve colaborar com a confiabilidade. Coerentemente com essa posição, optou-se pelo uso do coeficiente alfa de Cronbach, por meio do qual foram obtidos muito bons resultados e constatado não haver necessidade de aumentar a extensão do teste ou aumentar a amostra para melhorar o alfa. Na tabela 7 é possível observar quais itens podem melhorar a confiabilidade do instrumento se eliminados.

3.7. Etapa 7 — Testando a validade de construto

A validade do instrumento de medida diz respeito à congruência com a propriedade medida dos objetos e não com a exatidão com que a mensuração é feita. Essa afirmação de Pasquali (1997) tem uma posição contrária à prática comum de agrupar instintivamente vários itens e verificar *a posteriori* o que eles estão medindo, o que o autor denomina de “pesca de hipótese”.

A análise fatorial foi utilizada para estimar a validade das medidas empíricas e será considerada uma análise exploratória, uma vez que os fatores eram inicialmente desconhecidos. Compreende duas etapas principais da análise fatorial: a análise da matriz de correlação e a extração dos fatores.

- **Etapa 7a** — Matriz de teste de correlação para testar as correlações de todas as variáveis. Para os itens omissos, optou-se por sua inclusão pela média.
- **Etapa 7b** — Nesta etapa ocorre a extração dos fatores na análise fatorial. Pode-se observar os seus resultados na tabela 8.

Tabela 7
Confiabilidade (Alfa de Cronbach)

Item	Média de Escala se Item Deletado	Variância de Escala se Item Deletado	Item Corrigido – Correlação Total (Pearson)	Correlação Múltipla ao Quadrado	Alfa se Item Deletado
C1	66,7958	38,0218	0,4587	0,3548	0,8414
C2	66,7676	38,3924	0,4200	0,3267	0,8434
C3	66,5493	38,6039	0,5125	0,3870	0,8396
C4	66,6690	36,8897	0,5208	0,5785	0,8380
C5	66,7676	35,3711	0,6514	0,6432	0,8299
C6	66,5352	39,2576	0,4025	0,2674	0,8441
C7	66,8451	38,6850	0,4263	0,3290	0,8430
C8	66,8944	38,1661	0,4035	0,4785	0,8445
C9	66,9930	39,5106	0,2727	0,2970	0,8512
C10	66,7394	36,7189	0,5009	0,3502	0,8393
C11	66,5493	37,6678	0,5704	0,4592	0,8362
C12	66,4789	38,4499	0,5307	0,4547	0,8388
C13	66,8662	38,3862	0,3347	0,2995	0,8493
C14	66,8944	37,7122	0,4597	0,4642	0,8414
C15	66,8451	36,1035	0,6056	0,6129	0,8330
C16	66,5563	38,6174	0,5216	0,4152	0,8393

Nota: Gerada pelo SPSS®.

Tabela 8
Variância Total Explicada

Fator	Eigenvalues Iniciais			Soma dos Quadrados das Cargas da Extração			Soma dos Quadrados das Cargas da Rotação		
	Total	% da Variância	% Acumulado	Total	% da Variância	% Acumulado	Total	% da Variância	% Acumulado
1	5,206	32,535	32,535	4,695	29,343	29,343	2,186	13,662	13,662
2	1,530	9,565	42,100	1,080	6,748	36,091	1,932	12,072	25,734
3	1,331	8,319	50,419	0,800	4,999	41,090	1,721	10,756	36,490
4	1,199	7,492	57,911	0,687	4,295	45,385	1,423	8,895	45,385
5	1,025	6,403	64,314						

Nota: Gerada pelo SPSS®.

Foram feitas várias simulações com o *software* SPSS®. Na primeira, adotou-se o critério de Kaiser, quando foram selecionados os fatores cujos *eigenvalues* são maiores que 1 e obteve-se cinco fatores. O quinto fator tinha um *eigenvalue* muito próximo de 1; assim, optou-se por sua eliminação, por estar no limite de aceitação. A segunda e a terceira tentativas foram feitas pelos pesquisadores com quatro e seis fatores previamente estabelecidos. Após as tentativas com quatro e seis fatores, verificou-se que os melhores resultados foram aqueles obti-

dos com o critério de quatro fatores. Diante desses resultados, foi escolhida a opção com quatro fatores, descritos adiante.

O resultado das matrizes de fatores rotacionadas está descrito na tabela 9, na qual se constata o agrupamento dos quatro fatores ou dimensões em função da saturação de cada item nos respectivos fatores. O primeiro fator seria composto pelos itens C1, C2, C3, C10, C11 e C12; o segundo seria composto pelos itens C4 e C5; o terceiro pelos itens C13, C14, C15 e C16; e o último fator pelos itens C6, C7, C8 e C9.

Tabela 9
Matriz Rotacionada

Fator	1	2	3	4
C2	0,618			
C11	0,588			
C12	0,551			
C1	0,521			
C3	0,452			
C10	0,424			0,307
C15	0,310	0,725		
C14		0,624		
C13		0,536		
C16		0,477		
C4			0,813	
C5			0,731	
C8				0,689
C9				0,579
C7				0,442
C6				0,303

Método de Extração: Fatoração do Eixo Principal.
Método de Rotação: Varimax com Normalização Kaiser.
Rotação convergiu em 6 iterações.

Nota: Gerada pelo SPSS®

A etapa de interpretação e denominação dos fatores, a partir do conjunto de 16 itens válidos e confiáveis, foi feita pelos próprios pesquisadores. Como resultado dessa atividade, descrevem-se a seguir essas quatro dimensões.

Fator 1 — Praticabilidade (C1, C2, C3, C10, C11, C12) — É a característica que facilita ou induz à utilização de sistemas. Ao analisar os itens que compõem esta dimensão, pode-se resumir os seus aspectos mais importantes: flexibilidade, fácil entendimento, clareza, disponibilidade, confiabilidade e atualização. O cálculo da praticabilidade foi feito pela soma do escore de cada item que compõe esta dimensão. Para eliminar os possíveis efeitos da variabilidade do número de questões em cada fator normalizou-se o fator, com a divisão do resultado pelo escore máximo possível de cada fator.

Fator 2 — Precisão (C4, C5) — É a maneira de efetuar os cálculos corretos e sem erros. Esta dimensão contempla a precisão dos cálculos e a satisfação do usuário com relação a essa precisão. Ambas as questões derivaram do instrumento original de Doll & Torkzadeh (1988) e Palvia (1996). O termo **precisão** foi testado isoladamente no teste piloto e mostrou-se pou-

co claro para os respondentes. Para alguns significava **necessário** e para outros **cálculo exato**. O segundo item é complementar ao primeiro, ou seja, tenta verificar a importância da precisão do aplicativo específico para a sua atividade. O cálculo da precisão foi feito de maneira análoga ao da dimensão Praticabilidade (fator 1).

O principal resultado deste trabalho é a identificação das quatro dimensões da Satisfação do Usuário (Praticabilidade, Precisão, Disponibilidade e Adequação da Informação) a partir de um instrumento desenvolvido para esse fim.

Fator 3 — Disponibilidade (C13, C14, C15, C16) — É a predisposição para oferecer a informação desejada e atualizada, no tempo certo, com suporte e treinamento. Os aspectos mais importantes desta dimensão foram: a todo instante, treinamento, suporte técnico e dentro do prazo. A questão C13 corresponde à atualização *real time*, termo esse pouco compreendido, conforme constatação feita na fase preliminar da construção deste instrumento. O mesmo aconteceu com o item C16, que corresponde à temporalidade (*timeliness*). Devido à não-familiaridade dos respondentes com os referidos termos, optou-se por descrever melhor o seu significado, em vez de utilizar os termos comumente encontrados nos instrumentos existentes. O cálculo da disponibilidade foi feito de maneira análoga ao das dimensões anteriores.

Fator 4 — Adequação da informação (C6, C7, C8, C9) — Corresponde à forma e ao conteúdo da informação sob a visão do usuário. Os aspectos mais relevantes nesta dimensão foram: completo, suficientes, relevantes e formato adequado. O termo saída (*output*), presente na maioria dos instrumentos, tem pouco significado para os sujeitos da amostra utilizada. Foi necessário explicar melhor o termo **saída** nos itens C7, C8 e C9. O cálculo da adequação da informação foi feito de maneira análoga ao das dimensões anteriores.

4. INTERPRETAÇÃO DOS RESULTADOS

Inicialmente, interpretou-se a consistência interna por meio do coeficiente alfa de Cronbach, com resultados satisfatórios. O valor de alfa de Cronbach foi acima de 0,84 e está dentro dos valores aceitáveis, segundo sugestão de Nunally (*apud* CHAU, 1999) e de Litwin (1995). Os valores aceitáveis segundo Nunally são: 0,7 para pesquisas preliminares, 0,8 para pesquisa básica e 0,9 para pesquisa aplicada. Para Litwin, os valores aceitáveis são aqueles superiores a 0,70. O valor de alfa de Cronbach dos itens mostra não ser necessário eliminar nenhum item, uma vez que isso não melhora o valor de alfa de Cronbach obtido (0,84) (tabela 7). Nessa etapa, foi confirmado o número de itens do questionário final e foram derivadas as

dimensões do instrumento. Não houve necessidade de eliminar nenhum item do questionário. Os 16 itens foram considerados relevantes pelos pesquisadores e confirmados por meio de testes estatísticos. Obtiveram-se as interpretações apontadas a seguir.

- A matriz de correlação gerou 14 valores acima de 0,40 e nenhum valor muito alto (o maior valor foi de 0,702). As correlações significantes para $p < 0,01$ correspondem a 88 das 120.
 - Adequação da amostra: KMO 0,8 (valor bem acima do mínimo recomendado que é de 0,5).
 - Teste de significância geral: Bartlett = 751 com sig. = 0,000.
- Os quatro fatores representam 45% das variâncias de todos os itens (tabela 8).
- A matriz de fatores (matriz F) possui altas saturações ($> 0,42$), com exceção do item C6. O critério inicialmente utilizado foi o de *eigenvalue* acima de 1 (critério de Kaiser), sendo que o fator com o *eigenvalue* igual a 1,02 foi, posteriormente, eliminado. Outros testes foram realizados com quatro e seis fatores e a melhor representação foi com quatro fatores. A análise visual feita pelo teste Scree (desenvolvido por Cattell) confirmou essa escolha. Embora o item C6 pudesse ser eliminado em função de sua baixa saturação ($< 0,4$) (tabela 9), considerou-se ser ele importante para mensurar a satisfação e, portanto, optou-se por mantê-lo no instrumento.

Com esses resultados satisfatórios, poder-se-ia afirmar que a satisfação do usuário teria quatro dimensões: praticabilidade, precisão, disponibilidade e adequação da informação.

5. CONCLUSÕES

O principal resultado deste trabalho é a identificação das quatro dimensões da Satisfação do Usuário (Praticabilidade, Precisão, Disponibilidade e Adequação da Informação) a partir de um instrumento desenvolvido para esse fim. Esse instrumento teve origem em diversos instrumentos válidos identificados na literatura e foi adaptado para o usuário brasileiro. O instrumento foi considerado válido por meio de testes empíricos de confiabilidade e validade de construto, e pode indicar um caminho na busca de padrões para a mensuração da satisfação do usuário. Embora a validade não possa ser estabelecida por apenas um único estudo, o instrumento pode dar origem a uma boa ferramenta para medir a satisfação do usuário.

Os resultados da pesquisa ainda sugerem que a Satisfação seja decomposta em quatro dimensões como base para uma ferramenta de auxílio na implementação de sistemas de informações. Poderá ser utilizada para medir a satisfação do usuário pós-implementação, como também para avaliar a satisfação pré-implementação, de tal modo que os eventuais ajustes necessários possam ser efetuados. Como resultado, serão obti-

das maiores chances de sucesso na implementação do sistema de informações. Pode-se destacar o estudo de Machado & Costa (2002) como uma aplicação real derivada deste trabalho por meio de um projeto para medir a satisfação do usuário de um Sistema de Informação na prescrição e na distribuição de medicamentos em um hospital.

Os gestores devem fazer uso de instrumentos para medir a satisfação do usuário e obter maiores êxitos com menores riscos na implementação de sistemas, uma vez que os recursos existentes são escassos.

A falta de instrumentos para a medida do sucesso da Satisfação do Usuário é uma realidade e essa *proxy* não pode ser considerada como única medida, mas é uma das *proxies* alternativas mais importantes. Outras devem ser pesquisadas para abranger novos aspectos e complementar a Satisfação do Usuário. Novas pesquisas de impacto na eficácia, na produtividade e na vantagem competitiva devem ser realizadas para fomentar o desenvolvimento de novos instrumentos.

Não se procurou obter a representação perfeita do instrumento e sim a representação válida do construto. Embora ele esteja validado empiricamente, a interpretação dos resultados obtidos com a aplicação desse instrumento deve ser feita com o cuidado e as limitações que qualquer instrumento da área requer.

Acredita-se que o presente instrumento possa ser melhorado, considerando-se os seguintes aspectos:

- motivação dos respondentes — foi baixa a receptividade dos respondentes ao questionário;
- assimetria da escala — o que acarretou uma concentração nos três pontos mais altos da escala;
- aumento da amostra — para confirmar os resultados.

Os gestores devem fazer uso de instrumentos para medir a satisfação do usuário e obter maiores êxitos com menores riscos na implementação de sistemas, uma vez que os recursos existentes são escassos. Para os pesquisadores, existe a possibilidade de investigar e propor novas práticas gerenciais.

Em função dos resultados não-satisfatórios obtidos nas implementações de Sistemas de Informações, sugere-se a busca de novas formas de mensuração do sucesso desses sistemas. A satisfação do usuário não é a única forma, porém acredita-se que poderá, em conjunto com outras métricas, dar suporte adequado às implementações de Sistemas de Informações com sucesso. O estudo aqui relatado é uma das primeiras tentativas de adaptação e expansão de um instrumento com **sabor local**, para a realidade brasileira, para medir a satisfação do usuário em relação a sistemas de informações e, portanto, ajustes são necessários.^u

- AMOROSO, D.L.; CHENEY, P.H. Testing a causal model of end-user application effectiveness. *Journal of Management Information System*, v.8, n.1, p.63-89, Summer 1991.
- BAILEY, J.; PEARSON, S.W. Development of a tool for measuring and analysing computer user satisfaction. *Management Science*, v.29, n.5, p.530-545, May 1983.
- CALDAS, M.P.; WOOD JR., T. How consultants can help organizations survive the ERP frenzy. Submetido ao *Managerial Consultation Division* of the Academy of Management to be considered for presentation at the annual meeting in Chicago, August 1999.
- CARMINES, E.G.; ZELLER, R. *Reliability and validity assessment*. Newbury Park, CA: Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-017, 1979.
- CHAU, P.Y.K. On the use of construct reliability in MIS research: a meta-analysis. *Information & Management*, v.35, p.217-227, 1999.
- DELONE, W.H.; MCLEAN, E.R. Information system success: the quest for the dependent variable. *Information Research*, v.3, n.1, p.60-95, Mar. 1992.
- DOLL, W.J.; TORKZADEH, G. The measurement of end-user computing satisfaction. *MIS Quarterly*, v.12, n.2, p.259-274, June 1988.
- _____. The measurement of end-user computing satisfaction: theoretical and methodological issues. *MIS Quarterly*, v.15, n.1, p.5-10, Mar. 1991.
- EIN-DOR, P.; SEGEV, E. Organizational context and MIS structure: some empirical evidence. *MIS Quarterly*, p.55-68, Sep. 1982.
- ETEZADI-AMOLI, J.; FARHOOMAND, A.F. On end-user computing satisfaction. *MIS Quarterly*, v.15, n.1, p.1-4, Mar. 1991.
- FREITAS, H.; OLIVEIRA, M.; SACCOL, A.Z.; MOSCAROLA, J. O método de pesquisa survey. *Revista de Administração da Universidade de São Paulo (RAUSP)*, v.35, n.3, p.105-112, jul./set. 2000.
- GALLETA, D.F.; LEDERER, A.L. Some cautions on the measurement of user information satisfaction. *Decision Science*, v.20, n.3, p.419-438, 1989.
- GELDERMAN, M. The relation between user satisfaction, usage of information system and performance. *Information & Management*, v.34, p.11-18, 1998.
- HOFSTEDE, G. *Culture's consequence — international differences in work — related values*. Beverly Hills, CA: Sage, 1980.
- HWANG, M.I.; THORN, R.G. The effect of user engagement on system success: a meta-analytical integration of research findings. *Information & Management*, v.35, p.229-236, 1999.
- IGBARIA, M.; NACHMAN, S.A. Correlates of user satisfaction with end user computing. *Information & Management*, v.19, p.73-82, 1990.
- IVES, B.; OLSON, M.H. User involvement and MIS success. A review of research. *Management Science*, v.30, n.5, p.586-603, May 1984.
- IVES, B.; OLSON, M.H.; BAROUDI, J.B. The measurement of user information satisfaction. *Communication of the ACM*, v.26, n.10, p.785-793, Oct. 1983.
- JOSHI, K. The measurement of fairness or equity perceptions of management information systems users. *MIS Quarterly*, p.343-358, Sep. 1989.
- KERLINGER, F.N. *Metodologia da pesquisa em Ciências Sociais — um tratamento conceitual*. São Paulo: EPU, 1980.
- KIM, C.; SUH, K.; LEE, J. Utilization and user satisfaction in end-user computing: a task contingent model. *Information Resources Management Journal*, v.4, n.11, p.11-24, Fall 1998.
- KIM, K.K. Task characteristics, decentralization, and the success of hospital information systems. *Information & Management*, v.19, n.2, p.83-93, 1990.
- LITWIN, M.S. *How to measure survey reliability and validity*. Thousand Oaks, CA: Sage, 1995. Survey Kit 7.
- MACHADO, R.O.; COSTA, A.L. Measuring the end-user's satisfaction as an important component for the success of the information systems for drug prescription and distribution. In: BALAS, Florida, USA, 20-23 Mar. 2002.
- MATHIESON, K.; RYAN, T. The effect of definitional variations on user's evaluations of information systems. *Data Base*, p.37-48, May 1994.
- MELONE, N.P. A theoretical assessment of the user-satisfaction construct in information system research. *Management Science*, v.36, n.1, p.76-91, Jan. 1990.
- PALVIA, P.C. A model and instrument for measuring small business user satisfaction with information technology. *Information*

ABSTRACT

Development of an instrument to measure the satisfaction of information systems users

The objective of this study is to measure end-user's satisfaction as a contribution to the success of Information Systems (IS). An instrument was developed to measure end-user's satisfaction related to specific computer application selected by respondents. A sample of 143 responses from MBA end-users from School of Economics, Business and Accounting — University of São Paulo — Brazil was used and the results point to a 16-item instrument (five-point Likert-type scale) considered reliable with Cronbach's alpha 0.84 and validated by factor analysis and confirmed by multidimensional analysis of a variety of applications. Four components of end-user's satisfaction have been identified: practicability, availability, precision and information adequacy. This research highlights the importance of user's satisfaction for the success of an Information System. It also suggests that user's satisfaction, when decomposed into its 4 detailed components, could be the basis for a diagnostic tool in the implementation of the information system, thus improving the chances of its success.

Uniterms: user satisfaction, information system.

RESUMEN

Desarrollo de un instrumento para mensurar la satisfacción del usuario de sistemas de informaciones

Este trabajo tiene por objetivo desarrollar un instrumento para mensurar, empíricamente, la satisfacción del usuario de Sistemas de Informaciones. Para eso, se desarrolló un instrumento para medir la satisfacción del usuario en lo que se refiere a aplicativos específicos, elegibles por el mismo encuestado, a través del método *survey*. Se utilizó una muestra de 143 respuestas de participantes del MBA de la Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto (FEARP) de la Universidade de São Paulo, y el resultado apunta un instrumento (de 5 puntos en la escala Likert), cuyo coeficiente de consistencia interna, medido a través del coeficiente alfa de Cronbach, fue de 0.84, y cuya validación empírica se ha hecho por análisis factorial, confirmada, después, por el análisis multidimensional en un conjunto variado de aplicativos. Se identificaron cuatro componentes de la satisfacción del usuario: facilidad, disponibilidad, precisión y adecuación de la información. Esta encuesta sugirió, además, que la satisfacción del usuario cuando decompuesta en cuatro componentes, puede ser la base de una herramienta de diagnóstico en la implementación de sistema de información, lo que aumenta sus chances de éxito.

Palabras clave: satisfacción del usuario, sistemas de informaciones.

**CADERNO DE PESQUISAS EM ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO DA
FEA/USP**

O Caderno de Pesquisas em Administração é uma publicação trimestral para divulgação de trabalhos de natureza acadêmica, com temas relacionados à Administração. Nossos principais colaboradores são alunos e professores de Programas de Pós-graduação em Administração e alunos da graduação com orientação e co-autoria de professores. Todos os trabalhos publicados são avaliados em relação ao conteúdo e à metodologia.

Maiores informações: Editoria do Caderno de Pesquisas em Administração
E-mail: cpa@edu.usp.br
<http://www.ead.fea.usp.br/Cad-pesq>

