



ENSINO E PESQUISA EM
ADMINISTRAÇÃO

A INFLUÊNCIA DA ORDEM DAS QUESTÕES NOS RESULTADOS DE PESQUISAS SURVEYS

THE INFLUENCE OF ORDER OF THE ISSUES IN THE RESULTS OF RESEARCH SURVEYS

Celso Augusto de Matos
Doutor em Administração pela UFRGS
Universidade do Vale do Rio dos Sinos – UNISINOS

Guilherme Trez
Doutor em Administração pela UFRGS
Universidade do Vale do Rio dos Sinos – UNISINOS

Data de submissão: 20 ago. 2008 . **Data de aprovação:** 10 nov. 2011 . **Sistema de avaliação:** Double blind review. . Universidade FUMEC / FACE . Prof. Dr. Henrique Cordeiro Martins . Prof. Dr. Cid Gonçalves Filho.

RESUMO

A pesquisa *survey* é frequentemente utilizada nas pesquisas em Administração, sendo geralmente conduzida por meio de questionários autopreenchidos. Porém, alguns vieses podem prejudicar a qualidade dos resultados e da interpretação. O objetivo deste artigo foi avaliar os efeitos de ordem dos itens de um questionário nos resultados de uma *survey*. Para tanto, um questionário em quatro diferentes versões foi testado em uma amostra de 216 usuários de telefone celular. Os dados foram analisados por meio de testes de correlação, diferença de médias, análise fatorial e regressão múltipla. Os resultados mostraram não haver influência de ordem quando o respondente escolhia um subconjunto de itens dentre o conjunto total. Porém, o efeito ordem se mostrou significativo tanto na relação dos itens entre si quanto deles com um item geral, posicionado no final da escala. Por fim, esses resultados mostram a importância de se considerar um possível efeito ordem ao se planejar pesquisas *surveys*.

PALAVRAS-CHAVE

Efeito ordem. Questionários. Survey. Viés.

ABSTRACT

Surveys are a common method in the Management research, with investigations often based on self-fulfilled questionnaires. This procedure, however, might produce biases that can distort the research findings and interpretations. Hence, the aim of this research was to investigate the order effects of questionnaire items in survey research. A questionnaire was elaborated with four different versions and tested in a sample of 216 users of cell phone. Tests of correlation, mean difference, factor analysis, and multiple regressions were used in the data-analysis process. Findings suggested that there is no order effect when respondents are asked to choose the most relevant items among the set of items. However, an order effect was present and significant when the respondents answered the scale items and after that a final item evaluating the overall opinion about the scale's subject. Finally, these findings show the relevance of considering a possible order effects in survey research.

KEYWORDS

Order effects. Questionnaires. Survey. Bias.

INTRODUÇÃO

O objetivo desse trabalho é avaliar o efeito que a ordem dos itens de um questionário pode ter nos resultados e nas interpretações de uma pesquisa. O trabalho se baseia, principalmente, em dois tipos de situações: (i) uma em que é pedido, ao respondente, após avaliar os itens de um determinado constructo (qualidade, satisfação, etc.), para responder uma questão geral, de avaliação global daquele constructo; (ii) outra situação é quando o respondente

deve escolher alguns atributos dentre um conjunto total de considerações. Essa última situação foi analisada também por Carvalho e Leite (1997). Porém, esses autores deram foco na classificação de alguns atributos mais importantes dentre o conjunto total apresentado, mas não investigaram se a **ordem** influencia na **resposta a questões gerais**, usadas como avaliação global do constructo pesquisado. Assim, o presente artigo revê (replica) a questão da ordem, ao se escolherem alguns atributos/itens mais importantes, dentre o total de itens, e

amplia essa discussão para um possível efeito de ordem, quando se respondem questões gerais.

Inicialmente, é feita uma breve revisão da literatura sobre os conceitos de efeito ordem em *surveys*. Em seguida, são descritos os procedimentos utilizados para a coleta e análise de dados. Esses dados coletados foram analisados por meio de quatro técnicas diferentes, a fim de se permitir uma comparação dos resultados e de se obter maior embasamento para as conclusões. As técnicas incluíram testes de correlação, diferenças de médias, análise fatorial e regressão múltipla. A partir de então, discutem-se os resultados e sua relação com os estudos anteriores revisados. Finalmente, são apresentadas as contribuições e limitações deste estudo e novas possibilidades de pesquisa são delineadas.

O EFEITO ORDEM EM SURVEYS

É muito comum, em pesquisas das Ciências Sociais, o uso de instrumentos de coleta de dados desenvolvidos a partir de escalas de mensuração, que apresentam listas de indicadores para representar comportamentos, atitudes e percepções dos entrevistados. Nessas pesquisas, o esforço dos pesquisadores geralmente é direcionado para o controle de variáveis que possam causar distorções entre o resultado da mensuração e a verdadeira expressão da realidade.

Entre as diversas formas de controle do viés de resultados de pesquisas, estão os cuidados com a elaboração dos instrumentos de coleta (PRESSER *et al.*, 2004). Na elaboração dos questionários, a forma como as perguntas de múltipla escolha, característica de muitas escalas de avaliação de percepção

e comportamento, são apresentadas aos respondentes, pode afetar seu comportamento de resposta. Assim, os resultados de um estudo podem ser distorcidos a partir da sequência em que as perguntas são apresentadas, chamado efeito ordem, e do tamanho do questionário utilizado (BLUNCH, 1984).

As distorções resultantes do efeito ordem são discutidas com maior ênfase, desde que o problema foi apresentado em livros de pesquisa de mercado, cujo pioneiro foi *The Art of Asking Questions* (PAYNE, 1951), como citado em Blunch (1984). Segundo Schwarz (2003), o efeito ordem não se apresenta como um resultado padrão que possa ser tratado posteriormente em qualquer estudo: ele varia de acordo com o tema de pesquisa, com o público pesquisado e, principalmente, com a forma de aplicação dos questionários. Em pesquisas autopreenchidas, por exemplo, existe a tendência de o respondente ser mais criterioso com as primeiras respostas fornecidas; em outras pesquisas, nas quais se utilizam entrevistadores, por outro lado, existe a tendência de o entrevistado utilizar-se mais da reflexão referente à resposta anteriormente fornecida.

A existência de viés resultante do efeito ordem é explicada por Schwarz e Oyserman (2001), em três perspectivas. Como primeira causa, os autores destacam a fadiga na análise das questões em questionários autopreenchidos. Para os autores, os respondentes tendem a ser muito mais criteriosos ao responderem as primeiras perguntas, posteriormente, e, gradativamente, a fadiga faz com que menos elementos sejam considerados para cada resposta. A

segunda forma apresentada pelos autores ocorre em questionários aplicados por entrevistadores, pessoalmente ou por telefone. Nesse caso, os entrevistados não têm a possibilidade de ler as questões que lhes são apresentadas, sendo forçados a prestar mais atenção no que é dito no transcorrer da entrevista. O efeito da recenticidade faz com que os entrevistados elaborem suas respostas utilizando mais a ponderação utilizada para as respostas imediatamente anteriores. Há ainda uma terceira causa de viés, relacionada ao efeito ordem, que diz respeito a uma percepção de suficiência do entrevistado. Alguns entrevistados, em certo momento da pesquisa, passam a considerar que já responderam o suficiente e, a partir de então, deixam de responder ou respondem as questões formuladas sem a necessária ponderação.

Um estudo realizado no Brasil por Carvalho e Leite (1997) comparou a relação entre a ordem dos itens da escala SERVQUAL e as respostas dos pesquisados, em uma relação de fatores importantes. Os resultados mostraram não haver correlação entre as respostas de diferentes questionários, o que permitiu aos autores concluir pela ausência do viés de resposta na identificação de atributos importantes. Dessa forma, independentemente da ordem de apresentação dos atributos, a identificação da importância relativa dos mesmos permanece inalterada. Os autores ainda destacam outras formas de se avaliar a existência de viés, como os testes de comparação de médias e a análise fatorial, porém não desenvolveram tais análises.

Deve-se considerar ainda que a identificação dos atributos importantes, de maneira direta, ou seja, pela manifestação do entrevistado, tem dado lugar a alternativas de análise estatística, que permitem a redução do tamanho dos questionários. Autores como Rossi e Slongo (1998), por exemplo, sugeriram a utilização de perguntas gerais, após um conjunto de indicadores e a análise de regressão múltipla, como forma de se identificar a importância relativa de cada indicador na formação da avaliação geral, em um artigo que propõe um método brasileiro de pesquisa de satisfação de clientes.

Como essa forma de construção de questionários tem se tornado mais usual, é importante que se compreenda em que medida a sequência de apresentação dos itens pode causar viés na análise dos resultados de uma regressão múltipla. Essa é uma abordagem não mencionada no estudo de Carvalho e Leite (1997) e que deve ser considerada como alvo de testes do efeito ordem, principalmente devido à disseminação do uso da regressão múltipla, nas pesquisas em Administração.

Conforme defendeu Schwarz (2003, p. 593), deve-se ter a compreensão do impacto do efeito ordem em pesquisas e, assim, adotar o adequado procedimento metodológico, pois "nenhuma mágica estatística pode compensar esses problemas depois do fato".

MÉTODO

Foi realizado um *survey* com 216 usuários de telefone celular. Os dados foram coletados junto a alunos de graduação de uma instituição de ensino

superior, do sul do país. O questionário aplicado era composto de três partes:

(i) Variáveis demográficas: operadora, tipo de plano (pré-pago ou pós-pago), sexo, idade e gasto médio mensal;

(ii) Itens para avaliar a Qualidade dos Serviços prestados pela operadora: foi utilizada uma versão adaptada da escala SERVQUAL (PARASURAMAN; ZEITHAML; BERRY, 1988). Foram mantidos os 22 itens, mas utilizada apenas uma coluna,

com extremos variando de 1 (muito pior que o esperado) a 5 (muito melhor que o esperado), a fim de simplificar a resposta pelos participantes. As 22 questões buscam representar as dimensões teóricas do constructo qualidade em serviços: tangíveis, presteza, confiabilidade, garantia e empatia. A seguir, são apresentados, como exemplo, os itens referentes ao fator tangíveis.

QUADRO 1 – Forma de apresentação dos atributos no questionário

	Muito pior que o esperado	←————→				Muito melhor que o esperado
1. A operadora tem equipamentos modernos.	1	2	3	4	5	
2. A operadora tem instalações físicas com visual agradável.	1	2	3	4	5	
3. Os funcionários da operadora têm boa apresentação pessoal.	1	2	3	4	5	
4. Os materiais de divulgação da operadora têm visual moderno, agradável e inteligíveis.	1	2	3	4	5	
De 0 a 10, que nota geral você daria para a aparência das instalações físicas, dos equipamentos, do pessoal e dos materiais de divulgação da operadora?	NOTA:					

Essas dimensões foram separadas por blocos, no questionário, totalizando cinco blocos com 22 questões. Foi ainda incluída, ao final de cada bloco, uma questão geral, que tinha o objetivo de avaliar em que medida a ordenação dos itens dentro do bloco afetaria a nota geral do respectivo bloco. Ao final da escala de qualidade, havia uma questão que pedia ao participante para dar uma nota global para a qualidade dos

serviços da operadora, numa escala de 0 a 10. Em seguida, uma questão pedia ao respondente que enumerasse os seis atributos mais importantes, dentre os 22 apresentados até então, de forma idêntica ao usado por Carvalho e Leite (1997):

Das 22 questões apresentadas acima, quais delas são as mais importantes, na sua opinião? Indique abaixo com o número da questão:

Mais im-
portante (___) (___) (___) (___) (___) (___) Menos
importante

(iii) Itens para avaliar a Satisfação com a operadora de telefonia móvel: foi utilizada uma escala com 16 itens, oriunda de uma pesquisa prévia dos autores. Os respondentes poderiam optar por uma posição numa escala de 1 a 5, variando de muito insatisfeito a muito satisfeito. Ao final dessa escala, também havia uma questão para avaliar a satisfação geral com a operadora, numa escala de 0 a 10.

Dessa forma, os fatores combinados para a escala de qualidade foram: ordem dos blocos (2) e ordem dos itens dentro de cada bloco (2), totalizando em 4 combinações (ver QUADRO 2). Já na escala de satisfação, apenas duas variações foram feitas, em termos de ordem crescente ou decrescente dos 16 atributos.

Ressalta-se que, como foram 5 blocos, o terceiro se manteve nessa posição em todos os tipos. Essa decisão foi intencional, já que tal bloco representava uma dimensão tida como mais importante ao se avaliar a qualidade de serviços: a confiabilidade. Caso esse bloco viesse em uma posição inicial, em um tipo de questionário, e final, em outro, e houvesse relação entre as suas respostas e a resposta de qualidade geral, por exemplo, não seria possível separar esses dois efeitos: i) o fato de a dimensão ser mesmo a mais importante, como pesquisas anteriores têm mostrado; e ii) o fato de tal dimensão estar posicionada no início ou final do questionário. Essa discussão vai ser retomada ao se analisarem, por meio da regressão, os efeitos das notas gerais de cada bloco na nota geral da qualidade.

QUADRO 2 – Ordem das questões de Qualidade nos 4 tipos de questionários

Tipo 1	Tipo 2	Tipo 3	Tipo 4
t1	t4	e1	e5
t2	t3	e2	e4
t3	t2	e3	e3
t4	t1	e4	e2
qg_tangíveis	qg_tangíveis	e5	e1
pr1	pr4	qg_empatia	qg_empatia
pr2	pr3	g1	g4
pr3	pr2	g2	g3
pr4	pr1	g3	g2
qg_presteza	qg_presteza	g4	g1
c1	c5	qg_garantia	qg_garantia
c2	c4	c1	c5
c3	c3	c2	c4
c4	c2	c3	C3
c5	c1	c4	C2
qg_confiabilidade	qg_confiabilidade	c5	C1
g1	g4	qg_confiabilidade	qg_confiabilidade
g2	g3	pr1	pr4
g3	g2	pr2	pr3
g4	g1	pr3	pr2
qg_garantia	qg_garantia	pr4	pr1
e1	e5	qg_presteza	qg_presteza
e2	e4	t1	T4
e3	e3	t2	T3
e4	e2	t3	T2
e5	e1	t4	T1
qg_empatia	qg_empatia	qg_tangíveis	qg_tangíveis

ANÁLISE DOS DADOS

A fim de se atingir o objetivo, os dados foram analisados de quatro formas diferentes. Destas, três foram sugeridas por Carvalho e Leite (1997):

- 1) testes de correlação entre números de votos (Pearson), ou entre ordenações (Spearman, Kendall), das distintas versões de sequência dos atributos;
- 2) testes de diferenças de médias dos atributos entre os quatro tipos de questionários;

3) análise fatorial, a fim de se testar se as diferentes versões de questionários produzem estruturas de fatores subjacentes semelhantes.

Embora Carvalho e Leite (1997) tenham aplicado a primeira dessas sugestões, os mesmos ressaltam que as outras também são válidas. Nesta pesquisa, como o objetivo era avaliar também se a ordem dos atributos influencia a resposta que é dada para o item geral, optou-se por incluir um quarto tipo de análise:

4) testes de regressão múltipla para avaliar a importância relativa dos atributos entre as diferentes versões de questionários.

Teste de Correlação entre os Tipos

A primeira etapa desta pesquisa objetivou replicar os procedimentos adotados por Carvalho e Leite (1997). Por essa razão, utilizou-se a escala SERVQUAL (PARASURAMAN; ZEITHAML; BERRY, 1988), a forma de identificação dos atributos mais importantes e a análise de correlação proposta pelos mesmos.

A hipótese proposta em relação ao efeito ordem sugere que a sequência de apresentação das questões influencia a resposta dos entrevistados. Assim, foram comparados os atributos destacados como mais importantes para os quatro tipos de questionários desenvolvidos, a

fim de se verificar as diferenças entre os mesmos.

Utilizando-se dos mesmos procedimentos de Carvalho e Leite (1997), os atributos selecionados por cada entrevistado foram ponderados com um sistema de pontuação em que o mais importante recebeu peso 6, o segundo recebeu peso 5, e assim por diante, até o sexto, que recebeu peso 1. Assim, cada atributo foi pontuado de acordo com o número de indicações recebidas, multiplicado pela ponderação referente à resposta de cada entrevistado. Os resultados dessa análise foram quatro listas de 22 atributos, em que a cada um foi conferida uma pontuação específica (TAB. 1). As listas de atributos ponderados, originadas de cada um dos quatro tipos de questionários, foram comparadas, a fim de se verificar os indicadores de correlação entre as mesmas.

TABELA 1 – Pontuação Final de Cada Atributo em Cada Versão de Questionário

Atributos	V1	V2	V3	V4
1 A operadora tem equipamentos modernos.	39	18	42	40
2 A operadora tem instalações físicas com visual agradável.	13	8	6	5
3 Os funcionários da operadora têm boa apresentação pessoal.	11	10	2	30
4 Os materiais de divulgação da operadora têm visual moderno e agradável.	20	11	13	34
5 Os funcionários da operadora sempre atendem prontamente.	61	37	61	42
6 Os funcionários da operadora sempre demonstram boa vontade em atender os clientes.	48	46	37	49
7 Os funcionários da operadora estão sempre disponíveis para atender às solicitações dos clientes.	40	53	39	50
8 Os funcionários da operadora mantêm o cliente sempre informado sobre datas e prazos de entrega do serviço/produto.	56	25	46	96
9 A operadora sempre cumpre o que promete.	43	49	82	72
10 A operadora demonstra sincero interesse em resolver qualquer problema que surge.	70	46	58	53
11 A operadora faz as coisas corretamente da primeira vez.	26	46	33	44
12 A operadora mantém sempre corretos e atualizados os dados dos clientes.	19	33	17	28
13 A operadora sempre cumpre o que promete.	63	73	64	113
14 Como cliente, sinto-me seguro ao ser atendido pela operadora.	61	56	60	28
15 O comportamento dos funcionários da operadora gera confiança nos clientes.	20	26	23	20
16 Os funcionários da operadora são educados e corteses com os clientes.	41	28	22	36
17 Os funcionários da operadora têm competência para responder os clientes.	63	85	66	63
18 A operadora dá aos clientes atenção individual.	53	43	46	24
19 O horário de atendimento da operadora é conveniente para os clientes.	16	24	16	23
20 Os funcionários da operadora dão aos clientes um atendimento personalizado.	18	40	46	30
21 Os funcionários da operadora demonstram real interesse pelos clientes.	52	46	69	70
22 Os funcionários da operadora entendem as necessidades específicas dos clientes.	50	87	44	45

Nota – V1, V2, V3 e V4 se referem às quatro diferentes versões.

Na TAB. 2, são apresentados os resultados dos testes de correlação de Pearson e Spearman, calculados para os quatro tipos de questionários, em relação ao conjunto dos 22 atributos ponderados.

As correlações resultantes apresentaram valores significativos, o que indica a não influência da ordem das questões para a escolha dos atributos mais importantes.

TABELA 2 – Coeficientes de correlação de Pearson e Spearman

Correlações	TIPO1	TIPO2	TIPO3	TIPO4
TIPO1	1	0,670**	0,796**	0,658**
TIPO2	0,650**	1	0,727**	0,583*
TIPO3	0,804**	0,676**	1	0,652**
TIPO4	0,613**	0,479*	0,639**	1

Notas – Valores abaixo da diagonal se referem à correlação de Pearson e acima da diagonal à de Spearman; ** $p < 0,01$; * $p < 0,05$.

O teste de correlação de Pearson foi realizado, comparando-se os valores dos atributos ponderados para as quatro alternativas. O resultado demonstrou a inexistência de diferenças significativas relativas ao efeito ordem entre os questionários. Adicionalmente, como feito por Carvalho e Leite (1997), foi realizado o teste de correlação de Spearman. Esse teste permitiu a comparação do ordenamento das variáveis nas quatro alternativas de questionários. Novamente, os resultados demonstraram a inexistência de diferenças significativas entre o ordenamento dos quatro tipos.

Esses resultados demonstram que a ordem de apresentação dos atributos, em um questionário, não interferiu na identificação dos itens mais relevantes para o entrevistado. Dessa forma, a identificação direta de fatores importantes não foi afetada pelo efeito ordem.

Diferenças de médias

Nesse momento, são usados os testes univariados e multivariados de comparação de médias (ANOVA e MANOVA). As médias das questões de qualidade são apresentadas na TAB. 3.

TABELA 3 – Médias das questões de qualidade entre os quatro tipos de questionários

Questões	Médias					Diferença (sig) -
	tipo 1	tipo 2	tipo 3	tipo 4	total	
q1t1	3,74	3,71 ^a	3,88	4,11 ^a	3,86	0,058
q2t2	3,89	3,58 ^a	3,86	4,07 ^a	3,85	0,018
q3t3	3,68	3,40 ^a	3,80	3,82 ^a	3,68	0,051
q4t4	3,87	3,84 ^a	4,19	4,20 ^a	4,03	0,019
qg_tangíveis	7,85	7,70	7,56	8,18	7,82	0,151
q5pr1	3,04	2,76	2,95	2,96	2,93	0,636
q6pr2	3,00	2,90	3,18	3,09	3,05	0,544
q7pr3	2,89	2,69	3,03	3,00	2,91	0,311
q8pr4	2,92	2,61	2,75	3,02	2,83	0,269
qg_presteza	6,33	5,76	5,80	6,11	6,00	0,495
q9c1	3,28	2,94	2,80	2,93	2,99	0,174
q10c2	2,99 ^a	2,67	2,50 ^a	2,93	2,77	0,082
q11c3	3,04	2,86	2,84	2,85	2,90	0,757
q12c4	3,40	3,36	3,05	3,31	3,28	0,367
q13c5	2,76	2,52	2,42	2,66	2,59	0,439
qg_confiabilidade	6,32	5,84	5,48	6,08	5,93	0,257
q14g1	3,45 ^a	3,24	2,95 ^a	3,09	3,18	0,059
q15g2	3,30	3,31 ^a	2,82 ^a	3,09	3,13	0,094
q16g3	3,68	3,80	3,50	3,60	3,64	0,388
q17g4	3,26	3,06	2,80	3,11	3,06	0,130
qg_garantia	6,99 ^a	6,84	5,96 ^a	6,55	6,59	0,084
q18e1	3,49 ^{a,b}	3,18 ^c	2,48 ^{a,c}	2,85 ^b	3,00	0,000
q19e2	3,60	3,33	3,31	3,54	3,44	0,340
q20e3	3,15 ^a	3,12	2,72 ^a	2,74	2,93	0,049
q21e4	2,93	3,16 ^a	2,56 ^a	2,75	2,84	0,047
q22e5	3,06 ^a	3,22 ^{a,b}	2,47 ^{a,b}	2,84	2,89	0,002
qg_empatia	6,50	6,69*	5,58*	6,23	6,24	0,052
Qualidade Geral	7,31 ^a	7,04	6,23 ^a	7,08	6,90	0,029

Notas: Letras iguais (ex. ^a) indicam entre quais grupos as diferenças foram significativas.

Percebe-se que, em 14 das 28 variáveis, há pelo menos uma diferença significativa entre os grupos. É interessante notar, entretanto, que foi na questão 18 que apareceu um maior número de pares de grupos com diferenças, e que essa questão era a 18^o no tipo 1, 22^o no tipo 2, 1^o no tipo 3, e 5^o no tipo 4. Assim, houve diferença entre as médias dessa questão,

quando a mesma apareceu em 1^o lugar, no tipo 3, e no último, no tipo 2.

Interessante ainda ressaltar que nenhuma das diferenças significativas identificadas ocorreu entre os tipos 1 e 2, ou entre 3 e 4, e que esses eram muito similares entre si, na medida em que os blocos foram mantidos e variaram apenas as questões dentro dos blocos

(ver QUADRO 2). As diferenças que foram obtidas se deram entre os tipos 1 e 3, 2 e 4, e 2 e 3.

Para as questões de satisfação, a análise de variância também foi utilizada

e as médias das mesmas podem ser vistas na TAB. 4. Percebe-se, de forma geral, que não houve diferença entre as médias dessas questões, exceto para três delas: sat3 e sat9 (5%) e sat7 (10%).

TABELA 4 - Médias das questões de Satisfação entre os 2 tipos de questionários

Questões	Médias			Sig.
	Tipo 1 (AZ)	Tipo2 (ZA)	Total	
1. Clareza das informações sobre os planos e serviços.	3,18	3,00	3,09	0,203
2. Cobertura das ligações.	3,15	3,25	3,20	0,560
3. Conhecimento por parte dos(as) funcionários(as) das lojas da operadora quanto aos diferentes serviços.	3,49	3,04	3,26	0,000
4. Facilidade de contato com a Central de Atendimento da operadora.	2,94	2,78	2,86	0,343
5. Facilidade para acompanhar as despesas com as ligações.	3,15	2,91	3,03	0,128
6. Interesse dos(as) funcionários(as) das lojas da operadora em resolver os problemas e dúvidas apresentadas.	2,97	2,96	2,96	0,957
7. Opções de aparelhos disponíveis.	3,90	3,63	3,76	0,087
8. Preço das ligações.	2,45	2,57	2,51	0,460
9. Preço dos outros serviços oferecidos pela operadora.	2,47	2,92	2,70	0,002
10. Promoções de tarifas reduzidas.	2,90	2,94	2,92	0,837
11. Qualidade das ligações.	3,33	3,38	3,36	0,728
12. Qualidade dos serviços.	3,18	3,25	3,22	0,602
13. Rapidez na resposta pela Central de Atendimento.	2,69	2,67	2,68	0,891
14. Solução dada pela operadora aos problemas apresentados.	2,80	2,82	2,81	0,894
15. Tempo de espera para ser atendido nas lojas da operadora.	2,95	2,72	2,83	0,117
16. Variedade de serviços oferecidos pela operadora.	3,52	3,52	3,52	0,992
17. Satisfação Geral	6,66	6,36	6,51	0,275

Embora os resultados não mostrem diferenças significativas em todas as questões, as diferenças que foram encontradas parecem indicar que a ordem afeta as respostas, principalmente porque as diferenças maiores foram encontradas entre os tipos em que havia maior variação na ordem. Do contrário, se não houvesse diferença entre as médias nos diferentes tipos, poder-se-ia afirmar que

o tipo não importa e que os dados podem ser considerados em conjunto.

Análise Fatorial

Além das análises de correlação e de comparação de médias, outras formas de verificação da existência de efeito ordem são sugeridas na literatura. A análise fatorial é uma dessas formas, pois permite identificar a existência de

constructos latentes em uma escala de mensuração.

Nessa etapa, foi utilizada a análise fatorial exploratória, para a verificação da forma de agrupamento das variáveis apresentadas em diferentes formatos. Devido ao rigor da análise, que exige um número mínimo de 5 casos por variável analisada (HAIR *et al.*, 1998) somente foi possível a execução da mesma para as questões relativas à satisfação.

Foi realizada a análise fatorial, utilizando-se a extração dos componentes principais com a rotação ortogonal das variáveis e sem a definição prévia do número de fatores (constructos) extraídos. O mesmo procedimento foi utilizado para os dois tipos de questionários (Tipo 1 ordem AZ e Tipo 2 ordem ZA), a fim de se comparar os resultados obtidos.

Os resultados da análise fatorial foram avaliados, conforme proposto por Hair *et al.* (1998), a partir do teste de esfericidade de Bartlett, que demonstra a significância das correlações entre as variáveis da matriz, e da medida de

adequação da amostra Kaiser-Meyer-Olkin (KMO). Os resultados dessa análise são apresentados na TAB. 5.

Percebe-se, inicialmente, que os dois modelos analisados apresentam resultados gerais (KMO, Bartlett e significância) adequados. A análise da comunalidade dos atributos – variância total compartilhada por uma variável com as demais da análise, por outro lado, demonstra resultados diferentes. Para o Tipo 1, os resultados de comunalidade apresentam, em praticamente todos os atributos, valores superiores a 0,5, à exceção de sat3, cujo valor é 0,458, ou seja, próximo ao desejável. O Tipo 2, por outro lado, apresenta 4 indicadores com valores de comunalidade inferiores ao desejado, sendo ainda o resultado de sat7 (0,274) não indicado para o uso em análise fatorial. Em função da baixa comunalidade, e conseqüente baixa carga fatorial (inferior a 0,4), uma vez que há pouca variância compartilhada, a variável sat7 foi excluída dos fatores originados no Tipo 2. Esses resultados já demonstram que os modelos não são similares.

TABELA 5 – Análise Fatorial dos itens de Satisfação

Variáveis	Tipo 1 – Satisfação tipo AZ <i>KMO=0,804; Bartlett's test=554,43; sig.=0,000</i> <i>Variância explicada 63,4%</i>					Tipo 2 – Satisfação tipo ZA <i>KMO=0,870; Bartlett's test=704,88; sig.=0,000</i> <i>Variância explicada 69,1%</i>			
	Fator 1	Fator 2	Fator 3	Fator 4	Com.	Fator 1	Fator 2	Fator 3	Com.
sat1	0,593				0,579	0,580			0,488
sat2			0,712		0,625			0,742	0,630
sat3				0,495	0,458			0,637	0,604
sat4	0,659				0,554	0,702			0,611
sat5		0,632			0,566		0,486		0,409
sat6	0,661	0,370			0,604	0,561	0,475		0,563
sat7				0,654	0,537				0,274
sat8		0,766			0,780		0,885		0,801
sat9		0,714			0,732		0,724		0,675
sat10		0,798			0,748		0,694		0,638
sat11			0,839		0,776			0,875	0,794
sat12			0,738		0,693	0,666		0,428	0,655
sat13	0,795				0,741	0,804			0,710
sat14	0,842				0,757	0,782			0,696
sat15	0,764				0,626	0,650			0,465
sat16				0,769	0,677	0,632		0,414	0,573

Notas: Foram excluídas da análise as variáveis com cargas fatoriais inferiores a 0,4; Com.= comunalidade.

Em relação à composição dos fatores, o Tipo 1 originou quatro fatores, enquanto o Tipo 2 originou apenas três: com isso, a composição dos fatores é alterada. Alguns indicadores (sat3, sat12 e sat16) passam a constituir dimensões diferentes pela simples alteração da ordem em que as questões são apresentadas, o que confirma a existência de viés, relacionado ao efeito ordem, na análise fatorial deste estudo. Essas constatações reforçam a importância da consideração do efeito ordem, em estudos que utilizem a análise fatorial como processo de análise.

Regressão Múltipla

A técnica de regressão múltipla foi aplicada tanto na parte da qualidade quanto da satisfação. Na de qualidade, duas formas foram usadas, sendo a primeira a regressão das notas gerais de

cada bloco, na nota geral de qualidade; e, a segunda, a regressão dos itens de cada bloco, na nota geral do respectivo bloco. Já na parte de satisfação, a análise foi mais simples, fazendo-se a regressão dos itens na avaliação global de satisfação. A intenção é avaliar se os resultados são diferentes quando se consideram os diferentes tipos de questionários.

Análise dos Itens Gerais de Qualidade

No caso da qualidade, para a regressão das notas dos blocos na nota geral de qualidade, consideraram-se dois tipos de questionários: (i) o que ordenava os blocos em tangíveis, prestação, confiabilidade, garantia e empatia (tipos A e B); (ii) o que apresentava a ordem inversa: empatia, garantia, confiabilidade, prestação e tangíveis (tipos C e D).

TABELA 6 - Análise de Regressão – Itens gerais de Qualidade

Variáveis	Modelo 1 (tipos 1 e 2) tangíveis, prestação, confiabilidade, garantia, empatia					Modelo 2 (tipos 3 e 4) empatia, garantia, confiabilidade, prestação e tangíveis				
	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>
Constante	2,58	0,75		3,45	0,001	-0,26	0,47		-0,56	0,578
qg_tang	0,11	0,09	0,09	1,17	0,246	0,32	0,07	0,26	4,53	0,000
qg_prest	-0,02	0,08	-0,03	-0,26	0,796	0,10	0,07	0,12	1,49	0,139
qg_confi	0,36	0,07	0,52	5,05	0,000	0,25	0,07	0,29	3,69	0,000
qg_garan	-0,07	0,09	-0,09	-0,81	0,419	0,12	0,10	0,12	1,21	0,228
qg_empa	0,33	0,08	0,42	4,07	0,000	0,27	0,09	0,28	3,11	0,003

Notas: *B* = beta não padronizado; *SE* = erro-padrão; β = beta padronizado; *t* = teste *t*; Ajustes no modelo 1: R^2 ajust = 0,593; $F_{5,77} = 24,92$; $p < 0,000$. No modelo 2: R^2 ajust = 0,796; $F_{5,89} = 74,50$; $p < 0,000$.

Os resultados apresentados na TAB. 6 mostram que, no modelo 1 (questionários 1 e 2), os fatores significativos para explicar a qualidade geral foram confiabilidade e empatia. Já no modelo 2 (questionários 3 e 4), o fator tangíveis também foi relevante. Nota-se que, nos questionários tipo 3 e 4, o bloco referente a tangíveis estava na última posição.

O fator empatia foi significativo em ambos os modelos, tendo sido o último bloco em um dos modelos e o primeiro bloco no outro. Esse resultado é um indício importante do efeito do primeiro e do último, já discutidos na literatura sobre esse assunto (*primacy and recency effects*).

Já o fator confiabilidade se mostrou significativo em ambos os modelos e sua

posição não se alterou, nos diferentes tipos de questionários, conforme já explicado na seção de métodos. O resultado obtido reforça as pesquisas anteriores, que defendem a confiabilidade como dimensão mais importante da qualidade em serviços.

Análise dos itens em cada bloco

Nesse momento, as análises foram realizadas para testar o efeito da ordem dentro dos blocos, na nota geral do respectivo bloco. Assim, considerando a TAB. 7, foi realizada uma análise de regressão para cada cruzamento, entre tipo de questionário e bloco, totalizando 20 rodadas de análise de regressão. Em cada regressão, a variável dependente era o item geral do respectivo bloco.

TABELA 7 - Resultados da Análise de Regressão por blocos em Qualidade

Variáveis	Tipo 1 (AZ)			Tipo 2 (ZA)			Tipo 3 (AZ)			Tipo 4 (ZA)		
	β	t	p	β	t	p	β	t	p	β	t	p
Tangíveis												
Constante	-	2,30	0,026	-	2,88	0,006	-	,16	0,874	-	1,85	0,070
Q1T1	0,15	1,44	0,156	0,59	4,60	0,000	0,04	,31	0,759	0,24	1,68	0,099
Q2T2	0,37	3,10	0,003	0,16	1,30	0,200	0,27	1,60	0,117	0,17	1,16	0,252
Q3T3	-0,05	-0,43	0,670	0,00	-,04	0,971	0,14	,92	0,363	0,11	,78	0,440
Q4T4	0,50	4,08	0,000	0,18	1,38	0,175	0,38	2,62	0,012	0,26	1,87	0,067
Presteza												
Constante	-	1,826	0,075	-	2,24	0,031	-	-1,29	0,202	-	-1,48	0,146
Q5PR1	0,124	,897	0,375	0,46	3,35	0,002	0,22	1,88	0,066	0,42	3,75	0,000
Q6PR2	0,214	1,698	0,097	0,23	1,59	0,119	0,06	,45	0,651	0,14	1,31	0,196
Q7PR3	0,320	2,303	0,026	0,07	,48	0,635	0,45	3,29	0,002	0,39	3,98	0,000
Q8PR4	0,386	4,219	0,000	0,19	1,64	0,108	0,26	2,67	0,010	0,04	,51	0,615
Confiabilidade												
Constante	-	,63	0,533	-	1,09	0,281	-	-,15	0,881	-	,54	0,592
Q9C1	0,53	4,65	0,000	0,30	1,92	0,062	0,05	,43	0,672	0,48	4,51	0,000
Q10C2	0,31	2,93	0,005	0,24	1,53	0,135	0,30	2,67	0,011	0,22	2,21	0,032
Q11C3	0,10	,92	0,363	0,05	,46	0,648	0,15	1,32	0,194	0,04	,40	0,690
Q12C4	-0,03	-,33	0,740	-0,04	-,34	0,735	0,17	1,63	0,109	-0,02	-,24	0,810
Q13C5	0,08	,63	0,534	0,39	3,00	0,005	0,34	2,54	0,015	0,26	2,04	0,047
Garantia												
Constante	-	-,27	0,791	-	3,11	0,003	-	-1,31	0,196	-	1,24	0,221
Q14G1	0,15	1,49	0,143	0,44	3,22	0,002	0,12	1,22	0,230	0,49	3,60	0,001
Q15G2	0,43	3,34	0,002	0,28	2,06	0,045	0,55	5,03	0,000	0,32	1,95	0,056
Q16G3	0,10	1,19	0,242	-0,11	-,92	0,361	0,18	2,23	0,031	0,06	,48	0,633
Q17G4	0,35	3,62	0,001	0,26	2,09	0,043	0,20	2,06	0,045	-0,02	-,15	0,883
Empatia												
Constante		-,27	0,789		,89	0,379		,72	0,477		-,54	0,594
Q18E1	0,40	4,01	0,000	0,43	3,88	0,000	0,08	,66	0,510	0,42	3,56	0,001
Q19E2	0,09	1,02	0,315	0,05	,61	0,546	0,09	,90	0,373	0,22	2,19	0,033
Q20E3	0,09	,86	0,397	0,45	3,81	0,000	0,24	1,84	0,073	0,00	-,02	0,984
Q21E4	0,24	1,79	0,081	-0,04	-,37	0,712	,28	1,85	0,072	,14	1,04	0,304
Q22E5	0,26	2,03	0,049	0,12	1,24	0,221	,31	2,21	0,033	,33	2,78	0,008

Notas: β = beta padronizado; t = teste t ; De todas as análises, o R^2 ajustado variou de 0,34 a 0,79. A estatística F foi significativa em todos os casos (valor mínimo igual a 8 e máximo igual a 48, todos significativos a 5%).

Para se analisar a TAB. 7, é preciso antes lembrar que a ordem dos atributos é igual nos questionários tipo 1 e 3 (alfabética crescente), bem como entre 2 e 4 (alfabética decrescente). O motivo de os mesmos serem analisados separadamente se deve ao fato de haver diferença entre eles quanto à ordenação dos blocos (ver

QUADRO 2). Um modo simples de analisar essa tabela é comparar as células em destaque. Espera-se que as células, com o mesmo destaque, sejam significativas, e que as outras, **na mesma linha**, não o sejam.

De forma geral, é possível perceber, na TAB. 7, uma tendência de significância

do último item, nos questionários tipo 1 e 3, o que ocorreu nos blocos referentes a tangíveis, prestação, garantia e empatia. Isto é, independente da ordem em que o bloco aparecia no questionário, o último item tendia a ser um preditor significativo da variável nota geral daquele bloco.

Porém, uma evidência de efeito ordem existe se e somente se esses itens são significativos, quando aparecem em última posição, e **não** são significativos quando mudam de posição. Essa evidência esteve claramente presente nos blocos tangíveis e prestação, pois neles o último item foi significativo nos tipos 1 e 3, e não significativo nos tipos 2 e 4. Percebe-se ainda que, nesses dois blocos, nos questionários tipo 2 e 4, o primeiro item tendia a ser significativo; note-se que esse primeiro item apareceu na última posição, nesses dois tipos de questionários.

No bloco relativo a **confiabilidade**, nota-se uma influência tanto do primeiro quanto do último item: nos tipos 2 e 4 (quando q9c1 apareceu por último), houve tendência de significância, indício de efeito de posição. De forma similar, houve influência também da primeira posição, já que o item q13c5 foi significativo nos tipos 2 e 4 e, em ambos, esse item foi o primeiro da lista.

No bloco de **garantia**, o último item foi significativo nos tipos 1 e 3, quando efetivamente apareceu na última posição; porém, continuou sendo significativo no tipo 2, quando apareceu na primeira posição. Já avaliando o primeiro item dos tipos 2 e 4, quando apareceu por último, percebe-se que tal item foi significativo, reforçando a ideia de efeito da última posição, e não foi significativo quando apareceu, de fato, na primeira posição, nos questionários tipo 1 e 3.

No bloco **empatia**, houve também indício do efeito da última posição, embora não tão claro como nos blocos de tangíveis e de prestação, na medida em que o último item foi significativo nos tipos 1 e 3; porém, esse item (q22e5) foi significativo também no tipo 4, quando apareceu no início do questionário. De modo similar, quando se analisa o item q18e1, nota-se que o mesmo foi significativo nos tipos 2 e 4, quando veio em última posição, mas também o foi no tipo 1, quando apareceu no início da lista.

Análise dos itens em Satisfação

O efeito ordem, nas questões de satisfação, também foi avaliado por meio da análise de regressão, fazendo-se os 16 itens como variáveis independentes e, como variável dependente, a questão geral de satisfação, que estava posicionada logo após o final das questões relativas a esse constructo. Na TAB. 8 são apresentados os resultados obtidos nos dois modelos de questionários utilizados.

É possível observar, pela análise da TAB. 8, que os itens que foram significativos no modelo 1, quando os itens estiveram na ordem alfabética crescente; não se mantiveram significativos no modelo 2, quando os itens foram apresentados na ordem alfabética decrescente, e também que os itens significativos, no modelo 2, não o foram no modelo 1.

Um segundo tipo de regressão também foi usado, o de *stepwise*, a fim de se verificar se as mesmas variáveis entrariam no modelo em cada uma das duas amostras. Os resultados indicaram que, no modelo 1, as variáveis escolhidas foram sat6, sat11 e sat14, enquanto, no modelo 2, foram mais variáveis: sat12, sat10, sat3, sat16, sat13 e sat4.

TABELA 8 - Resultados da Análise de Regressão em Satisfação

Variáveis	Modelo 1 (ordem AZ); n= 72					Modelo 2 (ordem ZA); n= 83				
	B	SE	β	t	p	B	SE	β	t	p
Constante	1,33	,80		1,67	,100	-,50	,71		-,71	,481
SAT1	,11	,22	,07	,52	,605	,33	,19	,17	1,71	,092
SAT2	-,03	,15	-,02	-,19	,849	-,23	,16	-,13	-1,38	,173
SAT3	,15	,20	,09	,77	,447	,53	,26	,21	2,07	,043
SAT4	-,01	,16	-,01	-,06	,952	-,38	,20	-,22	-1,92	,059
SAT5	-,07	,16	-,05	-,45	,654	-,15	,17	-,08	-,88	,381
SAT6	,59	,22	,35	2,68	,010	,06	,24	,03	,26	,794
SAT7	-,19	,16	-,13	-1,23	,225	-,03	,16	-,02	-,20	,839
SAT8	-,11	,21	-,09	-,55	,581	,03	,18	,02	,17	,868
SAT9	,30	,21	,21	1,43	,158	,29	,22	,13	1,30	,199
SAT10	,12	,15	,09	,81	,421	,46	,18	,26	2,55	,013
SAT11	,44	,20	,31	2,17	,035	,24	,22	,12	1,09	,281
SAT12	,21	,24	,13	,87	,389	,38	,24	,19	1,61	,112
SAT13	-,30	,20	-,23	-1,50	,139	,46	,21	,26	2,20	,031
SAT14	,34	,20	,25	1,67	,101	,07	,22	,03	,31	,759
SAT15	-,02	,19	-,01	-,12	,908	-,20	,18	-,10	-1,11	,271
SAT16	,19	,20	,12	,98	,330	,34	,21	,16	1,62	,109

Notas: β = beta não padronizado; SE = erro padrão; b = beta padronizado; t = teste t; variável dependente: nota geral de satisfação; ajustes no modelo 1: R^2 ajust = 0,465; $F_{16,55}=4,86$; $p < 0,000$ e no modelo 2: R^2 ajust = 0,630; $F_{16,66}=9,72$; $p < 0,000$.

Esses resultados podem ser evidências do efeito ordem **se e somente se** essas diferenças não forem específicas de cada uma das subamostras. Para tanto, deve-se supor que essas subamostras são homogêneas em todas as outras variáveis. De fato, uma comparação de características como sexo, idade, gasto médio mensal, tipo de programa (pré-pago ou pós-pago) e marca da operadora, dentre os dois grupos, não mostrou diferenças significativas (foram feitos testes de qui-quadrado e de diferença de médias por análise de variância).

DISCUSSÃO DOS RESULTADOS E CONCLUSÕES

O uso de pesquisas do tipo *survey* é bastante comum nas ciências sociais, o que não é diferente para a disciplina de Administração. Na maioria dos casos,

as pesquisas são conduzidas por meio de questionários autoperenchidos pelos pesquisados. Pesquisas anteriores, como a de Duffy (2003), já mostraram o que se convencionou chamar de **efeito ordem**, ou seja, a influência da ordenação das questões nas respostas obtidas. Nesse contexto, dois vieses principais são abordados: quando um questionário é respondido por escrito, a primeira alternativa de uma dada questão tende a receber maior atenção (*primacy effect*) e, quando essa lista de opções é lida para o respondente, o mesmo tende a escolher a última opção (*recency effect*). No presente artigo, defende-se uma terceira possibilidade de viés: quando se usam questões gerais como avaliações globais de um determinado constructo, logo após o conjunto de itens do respectivo

constructo, em um questionário escrito: a resposta ao último item tende a ter maior influência na resposta que se dá ao item geral.

Assim, o objetivo do presente artigo foi avaliar os efeitos de ordem dos itens de um questionário nos resultados encontrados na pesquisa, particularmente nessas duas situações: uma em que o respondente, após avaliar os itens de um determinado constructo (qualidade; satisfação), é solicitado a responder uma questão geral, de avaliação global daquele constructo; outra situação é quando o respondente deve escolher alguns atributos dentre um conjunto total de considerações.

A análise dos resultados foi dividida em quatro partes, a fim de dar maior embasamento às interpretações e permitir comparações entre elas.

Na primeira parte, foi analisada a situação em que o respondente escolhe um subconjunto de itens dentre um conjunto total apresentado. Esse tipo de análise foi apresentado por Carvalho e Leite (1997). Esses autores pesquisaram usuários de serviços postais e 4 tipos de questionários foram usados, variando-se uma escala de qualidade em serviços de 39 itens. Na pesquisa aqui apresentada, com usuários dos serviços das operadoras de telefonia móvel, um conjunto total de 22 itens foi apresentado e, de forma similar ao estudo de Carvalho e Leite (1997), foi pedido ao respondente para escolher os seis itens mais importantes. Os resultados das correlações entre os pontos de cada item, em cada versão do questionário, mostraram que não havia efeito de ordem nessa situação, interpretação esta coerente com a apresentada por Carvalho e Leite (1997).

Na segunda parte, foram realizados os testes de diferença de médias, tanto para os itens de qualidade quanto de satisfação. Nos itens de qualidade, foram identificados vários itens com diferença significativa de média, entre os 4 tipos de questionários. Verificou-se ainda que essas diferenças foram mais frequentes quando se comparavam os tipos 1 versus 3, 2 versus 3 e 2 versus 4, que eram justamente os que possuíam maior variação na ordem. Assim, as comparações 1 versus 2 e 3 versus 4, quando a ordenação dos blocos era a mesma, não mostraram diferenças significativas. Para esses itens de qualidade, foi realizada ainda a versão multivariada do teste de diferença, considerando somente os itens de avaliação geral de cada bloco. Os resultados mostraram diferenças significativas nos itens referentes a garantia, empatia e qualidade geral, quando as suas médias foram comparadas entre os quatro tipos de questionários.

Já nos itens de satisfação, esse teste das diferenças de médias apontou menores diferenças (somente 3 itens significativos), o que pode ser relacionado ao fato de se ter apresentado somente dois tipos de ordens nesses itens (crescente e decrescente, sem a existência de blocos), o que reduziu as possibilidades de comparação entre dois grupos somente.

Na terceira parte, foi realizada uma análise fatorial da escala de satisfação para cada um dos dois tipos de questionários. Essa análise não foi realizada na escala de qualidade, em virtude das limitações de tamanho de amostra, pois, considerando o número mínimo de 5 casos por variáveis (HAIR *et al.*, 1998), seria necessária uma amostra de $22 \times 5 = 110$ casos em cada um dos tipos, sendo que foram obtidos

em torno de 50 casos em cada tipo de questionário.

Os resultados da análise fatorial, na escala de satisfação, mostraram diferenças entre os dois subgrupos: no primeiro deles, 4 fatores foram identificados, ao passo que, no segundo, somente 3 fatores emergiram. Esses resultados podem ser decorrentes da ordem dos itens ou de diferenças específicas dos respondentes. Porém, a fim de descartar essa segunda possibilidade, os questionários foram distribuídos aleatoriamente entre os respondentes e uma comparação de questões demográficas, como sexo, idade, gasto médio mensal, tipo de plano e operadora, mostrou não haver diferenças significativas, já que há sugestões, na literatura, de que o efeito ordem possa estar relacionado ao perfil de idade e escolaridade (KNÄUPER, 1999). Assim, espera-se que as diferenças entre os tipos de questionários se devam somente à ordenação dos itens.

Na quarta e última parte da análise, foi aplicada a análise de regressão dos itens da escala, com o item geral em três momentos: (i) itens gerais dos blocos com o item geral de qualidade; (ii) itens de cada bloco de qualidade com o item geral do respectivo bloco; e (iii) itens da escala de satisfação com o item geral de satisfação. **No primeiro momento**, verificou-se que a nota geral de tangíveis se tornou preditora significativa da qualidade geral no tipo 2, quando o bloco tangíveis foi posicionado como último bloco, não tendo sido significativa no tipo 1, quando aparecia em primeira posição. Esse resultado foi um indício de efeito ordem, na resposta do item geral. De forma similar, **no segundo momento**, as regressões específicas dentro de

cada bloco, entre os quatro tipos de questionários, também mostraram uma tendência de significância do último item, o que aconteceu de forma mais reforçada nos blocos de tangíveis e prestação, quando: (i) o último item foi significativo nos tipos 1 e 3 (quando apareceu por último) e não o foi nos tipos 2 e 4 (quando apareceu primeiro); e (ii) o primeiro item foi significativo nos tipos 2 e 4 (quando apareceu por último) e não o foi nos tipos 1 e 3 (quando apareceu em primeiro lugar), conforme destaques na TAB. 7. **No terceiro momento**, essa mesma análise de regressão foi aplicada também à escala de satisfação, considerando o item geral de satisfação como variável dependente, e os resultados mostraram diferenças de significância entre os preditores dos dois subgrupos comparados. Esses resultados também indicaram um possível efeito ordem, assumindo-se que não havia outras diferenças entre os grupos.

Algumas limitações, entretanto, devem ser ressaltadas: foi utilizada uma amostra não probabilística, de conveniência, composta por estudantes de graduação, embora se possa argumentar que amostras representativas são essenciais quando o objetivo é o de descrever (e inferir) as características da população pesquisada. Além disso, o uso de um subgrupo homogêneo da população total de proprietários de celular pode ser considerado um ponto forte para o objetivo do trabalho (comparar diferentes ordenações do questionário), já que as diferenças que surgiram entre os tipos de questionários podem ser mais seguramente atribuídas somente a esse fator e não a diferenças específicas entre os respondentes. Assim, ao se usar esses subgrupos mais homogêneos,

consegue-se reduzir a variância do erro (KERLINGER, 1980).

Outra limitação é que a escala SERVQUAL foi adaptada para os serviços de telefonia móvel, pelos pesquisadores, mantendo-se os 22 itens originais, e sem uma pesquisa qualitativa que poderia sugerir a inclusão de novos fatores relevantes ao contexto específico pesquisado.

Apesar dessas limitações, os resultados apresentados não só confirmam os que foram defendidos por Carvalho e Leite (1997), como também se estendem para o caso em que itens gerais são usados para capturar a avaliação global dos itens específicos de uma dada escala. Nesse último caso, os resultados permitem defender a hipótese de que a ordem dos itens influencia na avaliação geral. Portanto, os pesquisadores devem estar atentos a esse efeito e devem tentar utilizar diferentes tipos do instrumento de pesquisa, sempre que possível.

Os resultados obtidos com este estudo permitem sugerir, aos pesquisadores, uma maior atenção aos procedimentos metodológicos de elaboração de questionários e aplicação de pesquisas, no que diz respeito ao efeito ordem. Como diferenças importantes foram verificadas, nas análises de diferenças de médias, análise fatorial e de regressão múltipla, e, como essas são análises muito frequentes

em pesquisas tipo *survey*, pode-se sugerir que o efeito ordem possa distorcer resultados obtidos em muitos estudos.

A forma de tratamento do efeito ordem ainda é um tema que necessita ser mais explorado. Existem algumas sugestões na literatura (BLUNCH, 1984; DEMORANVILLE; BIENSTOCK, 2003; SCHWARZ, 2003) que precisam de maiores estudos para comprovação de seus resultados. São propostas as seguintes formas de prevenção ao viés do efeito ordem:

- Elaboração de diferentes tipos de questionários com o rodízio das questões;
- Inserção de perguntas que quebrem a rotina de resposta do pesquisado;
- Atenção ao treinamento de entrevistadores para a identificação de respostas não analisadas;
- Inserção de perguntas com escala inversa para avaliar a confiabilidade dos resultados obtidos.

São necessários novos estudos, que permitam uma maior compreensão dos efeitos decorrentes do viés de ordem, em pesquisas do tipo *survey*, pois até o momento sua existência é conhecida, mas não totalmente compreendida. É importante que se pesquisem, ainda, quais as formas de se verificar a existência do mesmo em estudos já realizados. ➤

REFERÊNCIAS

- BLUNCH, NJ. Position Bias in Multiple-Choice Questions. **Journal of Marketing Research**, [S. l.], v. 21, n. 2, p. 216-220, 1984.
- CARVALHO, Frederico A. de; LEITE, Valdecy F. A Ordem dos Atributos Afeta a Avaliação de Qualidade? Uma Investigação Empírica a partir da Versão mais Recente do Modelo SERVQUAL. **Revista de Administração Contemporânea**, [S. l.], v.1, n.1, p. 35-53, 1997.
- CARVALHO, Frederico A. de; LEITE, Valdecy F. Alternativas de Ordenação da Importância de Atributos da Qualidade de Serviços: um estudo exploratório sobre o efeito do tamanho do *choice set*. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO – ENANPAD, 22., 1998, Foz do Iguaçu. **Anais...** Rio de Janeiro: ANPAD, 1998.
- DEMORANVILLE, Carol W.; BIENSTOCK, Carol C. Question order effects in measuring service quality. **International Journal of Research in Marketing**, [S. l.], v. 20, p. 217-231, 2003.
- DUFFY, B. Response order effects - how do people read? **International Journal of Market Research**, [S. l.], v. 45, n. 1, p. 457-475, 2003.
- HAIR, Joseph F., Jr.; ANDERSON, Rolph E.; TATHAM, Joseph L.; BLACK, William C. **Multivariate Data Analysis**. 5th. ed. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- KERLINGER, Fred N. **Metodologia de pesquisa em Ciências Sociais: um tratamento conceitual**. São Paulo: EPU, 1980.
- KNÄUPER, Bärbel. The Impact of Age and Education on Response Order Effects in Attitude Measurement. **Public Opinion Quarterly**, [S. l.], v. 63, n. 3, p. 347-370, 1999.
- MALHOTRA, N. **Pesquisa de marketing: uma orientação aplicada**. Porto Alegre: Bookman, 2001.
- MARCHETTI, Renato; PRADO, Paulo. Avaliação da Satisfação do Consumidor Utilizando o PLS: Um Modelo Aplicado ao Setor Elétrico Brasileiro. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO – ENANPAD, 25., 2001, Campinas. **Anais...** Rio de Janeiro: ANPAD, 2001.
- PARASURAMAN, A.; ZEITHAML, V. A.; BERRY, L. L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. **Journal of Retailing**, [S. l.], v. 64, n. 1, p. 12-41, 1988.
- PRESSER, Stanley; COUPER, Mick P.; LESSLER, Judith T.; MARTIN, Elizabeth; MARTIN, Jean; ROTHGEB, Jennifer M.; SINGER, Eleanor. Methods for Testing and Evaluating Survey Questions. **Public Opinion Quarterly**, [S. l.], v. 68, n. 1, p. 109-130, 2004.
- ROSSI, Carlos A. V.; SLONGO, Luiz A. Pesquisa de Satisfação de Clientes: O Estado-da-Arte e Proposição de um Método Brasileiro. **Revista de Administração Contemporânea**, [S. l.], v. 2, n.1, p.101-125, 1998.
- SCHWARZ, Norbert. Self-Reports in Consumer Research: the challenge of comparing cohorts and cultures. **Journal of Consumer Research**, [S. l.], v. 29, n. 4, p. 588-594, 2003.
- SCHWARZ, Norbert; OYSERMAN, Daphna. Asking Questions about Behavior: Cognition, Communication, and Questionnaire Construction. **American Journal of Evaluation**, [S. l.], v. 22, n. 2, p.127-160, 2001.