



Administração: Ensino e Pesquisa

ISSN: 2177-6083

raep.journal@gmail.com

Associação Nacional dos Cursos de
Graduação em Administração
Brasil

TEIXEIRA CARNEIRO, TEIXEIRA CARNEIRO; DA ROCHA DIB, LUÍS ANTÔNIO
O USO DA INTERNET EM SURVEYS: OPORTUNIDADES E DESAFIOS

Administração: Ensino e Pesquisa, vol. 12, núm. 4, outubro-novembre, 2011, pp. 641-670

Associação Nacional dos Cursos de Graduação em Administração

Rio de Janeiro, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=533556770004>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal

Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

O USO DA INTERNET EM SURVEYS: OPORTUNIDADES E DESAFIOS

INTERNET-BASED SURVEYS: OPPORTUNITIES AND CHALLENGES

JORGE MANOEL TEIXEIRA CARNEIRO (*jorgemtc@iag.puc-rio.br*)

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO/IAG

LUÍS ANTÔNIO DA ROCHA DIB

COPPEAD/UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

RESUMO

Este artigo discute o uso da Internet nas pesquisas de levantamento de dados (*surveys*), tratando das questões que são exclusivas deste meio e daquelas comuns a outros ambientes de pesquisa, mas que apresentam particularidades quando no ambiente *online*. Inicialmente discutem-se quatro tipos de erros, que costumam estar presentes em *surveys* – erros de cobertura, amostragem, não resposta e mensuração – e quais as causas e manifestações específicas destes no ambiente *online*, bem como as formas de se lidar com eles. Em seguida, discute-se o impacto do formato do questionário sobre a qualidade das respostas. Questões éticas como a garantia de confidencialidade e de anonimato são abordadas. Este material poderá ser útil a pesquisadores brasileiros – que poderão tirar proveito das potencialidades do novo meio ao mesmo tempo em que evitam suas “armadilhas” – uma vez que reúne um conjunto de orientações conceituais e práticas para o uso de *web surveys* que, de outra forma, somente se encontram disponíveis em diversos trabalhos isolados ou em extensos textos.

Palavras chave: pesquisas *online*; *web surveys*; coleta de dados.

ABSTRACT

This paper addresses the use of the Internet in surveys, including issues specific to the medium, as well as those common to other research mediums that feature specific characteristics when online. The paper begins with the presentation of four types of errors that are frequently found in surveys: coverage, sampling, non-response, and measurement errors. The specificities of the Internet – in terms of the causes and manifestations of these errors and ways of dealing with them – are discussed. The impact of the online questionnaire's format on the quality of the answers is also addressed and ethical concerns, such as confidentiality and anonymity, are looked at. This paper can help Brazilian researchers to make the most of this new medium while avoiding its methodological pitfalls. Its main contribution lies in the fact that it consolidates several conceptual and practical recommendations that are otherwise only found scattered across multiple works or long texts.

Keywords: *online research; web surveys; data collection.*

INTRODUÇÃO

A metodologia de coleta de dados em levantamentos (*surveys*) tem se modificado nos últimos tempos. O ritmo da vida moderna e o rápido desenvolvimento de novas tecnologias tiveram grande impacto sobre as maneiras de conduzir pesquisas. As pesquisas por telefone se tornaram mais difíceis de administrar, devido ao uso crescente de aparelhos celulares e ao menor uso de telefones fixos. Em muitos países, já está se consolidando uma cultura de inclusão de números residenciais ou comerciais em listas do tipo “*do-not-call*”, que vedam seu uso por empresas de pesquisa ou para a venda de produtos ou serviços. A alternativa das pesquisas por *e-mail* também pode ser pouco efetiva, devido ao excesso de apelos de vendas feitos por este meio. Já o alcance da *World Wide Web* e o crescimento na disponibilidade de *softwares* específicos e de *web hosts* para questionários e bases de dados fazem do ambiente *online* um meio promissor para pesquisas.

Diversos cuidados são relevantes a todos ambientes de pesquisa e também são aplicáveis ao ambiente *online*, tais como motivação aos respondentes, criação de questões válidas e a definição de procedimentos apropriados de análise dos dados. Mas o ambiente *online* e suas

ferramentas permitem facilidades não disponíveis em outros métodos de coleta de dados, tais como interatividade e fluxo dinâmico das questões baseadas em respostas providas pelo usuário. Contudo, o uso da *Internet* como meio para coleta de dados apresenta também diversos desafios.

Albaum, Roster, Wiley, Rossiter e Smith, (2010) argumentam que, embora muito já tenha sido debatido sobre as vantagens e desvantagens do uso de *web surveys* em comparação com outros formatos tradicionais de condução de *surveys*, ainda há muito a se investigar sobre a multiplicidade de opções para a preparação e aplicação de questionários *online*. São as vantagens e os desafios de coleta de dados que sejam específicos da *Internet*, bem como as formas de se lidar com eles, que serão aqui abordados.

O presente estudo se constitui em uma revisão organizada da literatura sobre *web surveys* (também designadas por *online surveys*). A contribuição principal deste artigo é a consolidação da discussão sobre o uso da *Internet* na coleta de dados para estudos científicos, uma vez que tal literatura se encontra espalhada por múltiplos trabalhos ou exposta em extensos textos. Acredita-se, assim, que esteja sendo oferecido material útil a pesquisadores brasileiros interessados em se valer deste relativamente novo e promissor meio de pesquisa.

Os tópicos aqui abordados estão circunscritos ao ambiente de pesquisas científicas e não a outros tipos de pesquisa (por ex. para fins de entretenimento ou pesquisas abertas à participação livre e voluntária, COUPER, 2000), que também são frequentemente realizadas na *Internet*. Além disso, somente são tratadas pesquisas de cunho quantitativo (especificamente, *surveys*; vide BEST; KRUEGER, 2004.), não se tendo abordado as diversas formas de pesquisa qualitativa (por ex. fóruns, grupos de discussão etc.), que também podem ser conduzidas via *Internet*.

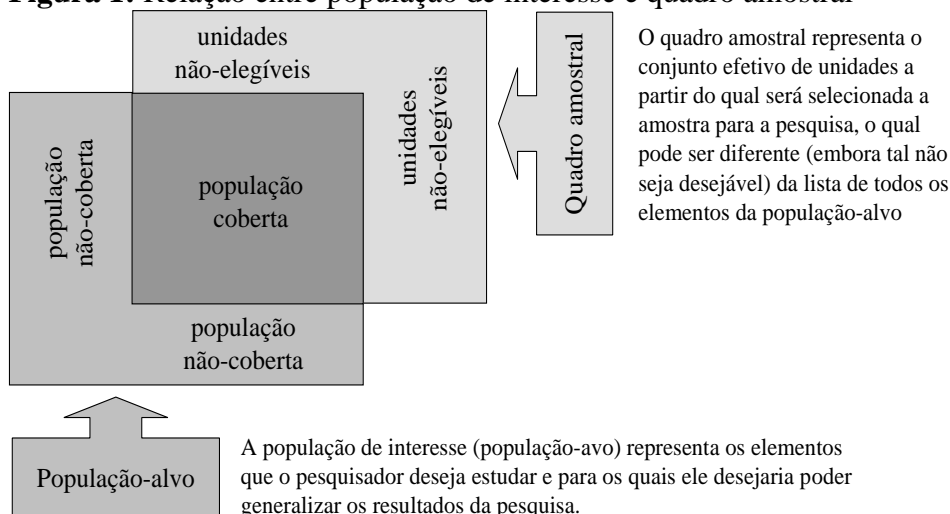
Este artigo está organizado em seis partes. Após a introdução, são apresentados os principais tipos de erros usualmente presentes em *surveys*, inclusive com o uso de meios tradicionais – erros de cobertura, de amostragem, de não resposta e de mensuração – com destaque para suas causas e as formas como tais erros se manifestam na *Internet*, bem como as estratégias para se lidar com eles nesse novo meio. Na terceira parte, são discutidas opções de formato do questionário *online* e seus efeitos sobre a qualidade dos dados obtidos. Em seguida, é apresentada uma discussão sobre ética em pesquisas *online*. Uma seção de considerações finais e o apêndice A com informações sobre ferramentas de pesquisa *online* fecham o artigo.

PRINCIPAIS TIPOS DE ERROS EM *SURVEYS* E SUAS PARTICULARIDADES NO AMBIENTE *ONLINE*

Pesquisas quantitativas estão sujeitas, de forma geral, a diversos tipos de erro, entre os quais os seguintes quatro: erro de cobertura, erro de amostragem, erro de não resposta e erro de mensuração. De forma a ser possível fazer-se uma generalização dos resultados de uma pesquisa para a população-alvo, é necessário que cada um dos erros a seguir detalhados seja mantido em nível razoavelmente baixo (DILLMAN; BOWKER, 2001). Nesta seção são discutidas as manifestações desses erros e o respectivo tratamento recomendado no caso particular de pesquisas *online*.

Antes da discussão sobre os quatro tipos de erro, é conveniente recordar-se as distinções entre população, quadro amostral e amostra. A população de interesse (população-alvo) de um estudo envolve todos aqueles elementos sobre os quais se deseja obter informações e para os quais seria desejável traçar inferências baseadas nos resultados da pesquisa, ou seja, é a coleção completa de todas as unidades de análise relevantes para a pesquisa em pauta (MALHOTRA, 2006).

Figura 1: Relação entre população de interesse e quadro amostral



Fonte: Enanoria (s/d).

Quadro amostral (*sampling frame* ou *frame population*) é a lista de potenciais participantes, que pode ou não coincidir com a população-alvo, de onde serão efetivamente selecionados elementos para compor a amostra do estudo; nas palavras de Enanoria (s/d), trata-se da definição

operacional da população-alvo. Amostra é o efetivo conjunto de elementos selecionados (com base no quadro amostral e não necessariamente na população, como seria desejado), sobre os quais se obtiveram os dados para um determinado estudo (MALHOTRA, 2006). A Figura 1 ilustra a relação entre população e quadro amostral.

Erro de cobertura (coverage error). Ocorre quando os membros da população não têm uma probabilidade conhecida e maior do que zero de serem selecionados para participação na pesquisa, o que significa que alguns membros da população não têm chance de ser selecionados, enquanto outros podem ter múltiplas chances (por ex. por aparecem múltiplas vezes na lista) e alguns membros do quadro amostral podem a rigor não atender aos requisitos da população-alvo, devendo, portanto, ser considerados inelegíveis para a pesquisa. Ou seja, erro de cobertura é o resultado de desajuste entre a população-alvo da pesquisa e o quadro amostral (Figura 1). O erro de cobertura resulta de uma combinação da proporção da população não coberta (e também aquela indevidamente coberta) pelo quadro amostral e da diferença entre membros cobertos e aqueles não cobertos (e também os inelegíveis), no que tange aos valores das variáveis relevantes para o estudo e sua distribuição estatística.

Erro de amostragem (sampling error). Resulta de se colherem dados somente junto a um subgrupo da população de interesse (na prática, do quadro amostral), ao invés de junto a todos os membros desta população. A questão então passa a ser o grau em que a amostra poderia ser considerada ou não representativa da população.

Erro de não resposta (nonresponse error). Ocorre devido ao fato de alguns membros selecionados para comporem a amostra não terem respondido à pesquisa (ou ao menos não terem respondido a todas as questões). O erro de não resposta é função da taxa de não resposta bem como da diferença entre respondentes *vs.* não respondentes em relação às suas respostas a variáveis focais da pesquisa. Neste caso é importante verificar se poderia ser esperado que os não respondentes teriam exibido um padrão de respostas que seria distinto do daqueles que efetivamente participaram da pesquisa. Se for julgado que o perfil dos não respondentes é diferente do dos respondentes em variáveis que poderiam ter relação com as variáveis de interesse da pesquisa, então o conjunto de respondentes não poderia ser considerado representativo da população-alvo.

Erro de mensuração (*measurement error*). É o resultado de respostas “incorretas” ou “imprecisas”, ou seja, desvios entre a resposta dada e o valor “verdadeiro” (HAIR; BLACK; BABIN; ANDERSON; TATHAM, 2006) e pode ser randômico ou sistemático. Usualmente, tal tipo de erro é decorrente de diferenças (entre pesquisador e respondente) de interpretação das perguntas ou do significado das opções de resposta, intervenção indevida do pesquisador ou do instrumento de pesquisa, viés de método (ou seja, o próprio instrumento de pesquisa induz ou influencia o tipo de resposta) ou deturpação intencional da resposta por parte do respondente.

A seguir, mostram-se como estes quatro tipos principais de erros se manifestam mais frequentemente em pesquisas *online*.

Erros de cobertura em pesquisas online

Erros de cobertura podem ocorrer porque alguns membros da população de interesse podem não ter acesso à *Internet* (resultando em um conjunto de unidades não coberto pela pesquisa, SUE; RITTER, 2007), enquanto outras pessoas que não seriam alvo da pesquisa podem acabar respondendo à mesma. Mesmo que todos os membros da população-alvo tenham acesso à *Internet*, ainda assim, erros de cobertura na *Internet* podem ocorrer em função de nem sempre ser possível obter uma lista de endereços (eletrônicos) de toda a população de interesse (COUPER, 2000), o que obriga o pesquisador a trabalhar com um quadro amostral (de onde serão, posteriormente, selecionados os elementos para compor a amostra), que incluiria apenas uma parte da população-alvo.

Ademais, haveria um erro de cobertura se fosse permitido a um dado respondente participar da pesquisa mais de uma vez. O envio de múltiplos questionários respondidos pelo mesmo respondente pode ocorrer de maneira acidental (por ex. quando o usuário aciona não intencionalmente o botão de <enviar> ou o programa não tem controle sobre o envio de questionários incompletos e o usuário resolve continuar respondendo e enviar novamente), inadvertida (quando o usuário, por demora do programa em enviar confirmação do envio ou por ausência desta, clica mais de uma vez na opção de <enviar>) ou maliciosa (quando o usuário tentar burlar o sistema) (STANTON; ROGELBERG, 2001).

Tais ocorrências de múltiplos questionários respondidos podem ser minimizadas por meio de orientação aos respondentes de que a pesquisa somente deve ser respondida uma única vez e o questionário enviado somente quando estiver completo, sendo permitido ao usuário salvar suas respostas parciais a qualquer momento para continuar depois. Além disso, se o usuário tentar enviar um questionário incompleto, ele

deve ser informado sobre as perguntas que ainda não foram respondidas, mas deve-lhe ser dada a possibilidade de confirmar a submissão do questionário ainda assim. Já para evitar tentativas maliciosas de múltiplas respostas, podem ser utilizados *cookies* (pequenos arquivos que são gravados no computador do usuário), de forma que um computador com o respectivo *cookie* da pesquisa já gravada não poderia ser utilizado para responder novamente à pesquisa (COUPER, 2000). No entanto, o mesmo usuário poderia se utilizar de outro computador ou poderia, se for razoavelmente proficiente na *Internet*, simplesmente apagar o arquivo do *cookie*. Nestes casos, o uso de senhas ou de identificações pessoais embutidas no endereço (*link*) da pesquisa (distribuídas previamente por *e-mail*, por ex.) pode ser a solução (COUPER, 2000). Vale ressaltar, ainda, que o uso de *cookies*, como forma de coibir múltiplas respostas, torna-se inadequado nos casos em que dois ou mais diferentes integrantes do quadro amostral poderiam compartilhar o mesmo computador (por ex. em escolas ou em algumas organizações).

Uma forma de evitar acessos múltiplos ou por elementos não autorizados a participar da pesquisa, diminuindo assim erros de cobertura, é enviar-se uma carta-convite via *e-mail* com o *link* para a página da pesquisa e informações de *login* e senha – uma tendência de pesquisas *online* mais recentes (por ex. DIB; ROCHA; SILVA, 2010).

Em pesquisas que empregam amostragem sistemática, onde um em cada n elementos da população (a rigor, do quadro amostral) é “convidado” a participar, uma técnica usualmente empregada é de interrupção/interceptação, onde o n -ésimo visitante a um dado *site* é apresentado à pesquisa (por meio de uma janela *pop-up*, por ex.) e assim por diante, de n em n visitantes ao *site* (BRADLEY, 1999). Há aqui um problema de definição do quadro amostral e, portanto, um possível erro de cobertura, pois somente estariam nele incluídos visitantes ao referido *site*, mas não necessariamente todos os membros da população-alvo. Além disso, ao contrário do processo de seleção sistemática em pesquisas tradicionais, em que o pesquisador pode usar de sua persuasão para incentivar a pessoa a participar, numa pesquisa na *Internet* é mais provável que somente aqueles usuários mais motivados com o tema venham efetivamente a participar (COUPER, 2000), o que tenderia a aumentar a taxa de não resposta e eventualmente resultar em viés de não resposta, dado o caráter voluntário da participação.

Por outro lado, quando a pesquisa envolve apenas membros de uma dada organização, então pode ser possível não apenas definir-se claramente a população de interesse, como também construir-se um

quadro amostral que reflita com bastante precisão a população-alvo, minimizando-se assim o erro de cobertura.

Erros de amostragem em pesquisas online

Em uma pesquisa na *Internet*, raramente é possível calcular a probabilidade de seleção de cada unidade da população-alvo. Em consequência, também não há condições para se calcular o erro de amostragem. Deve ser enfatizado que o simples fato de que pesquisa por meio da *Internet* permitir acesso (em tempo mais curto e com menos gastos) a um maior número de respondentes, não corresponde necessariamente a uma diminuição do erro de amostragem (COUPER, 2000), uma vez que, sem um mecanismo probabilístico de amostragem, não é possível calcular intervalos de confiança e fazer inferências para a população de interesse. Bradley (1999) sugere, inclusive, que o próprio horário de liberação de uma pesquisa (por ex. por meio do envio de *e-mails* para potenciais respondentes) pode ter influência sobre a composição da amostra.

Uma forma de se estimar os erros de amostragem em uma pesquisa *online* é a realização da mesma pesquisa com uma amostra selecionada por meios tradicionais, conforme sugerido por Simsek e Veiga (2001). O pesquisador poderia, então, comparar se há diferenças nos padrões de resposta (por ex. distribuição estatística, estrutura fatorial ou viés de não resposta). Se os padrões de resposta dos dois grupos forem similares, e desde que a amostra *online* tenha características probabilísticas, então se poderia argumentar que a amostra *online* seria razoavelmente representativa da população em geral. Mais ainda, se houver dados sobre o perfil “demográfico” da população, então pode ser realizada uma estratificação dos respondentes *online* (SINSEK; VEIGA, 2001), atribuindo-se pesos e proporções de repostas conforme a distribuição populacional (por ex. gênero, idade etc.).

Embora este mecanismo de controle aumente o esforço de pesquisa, pelo fato de se utilizarem dois métodos de coleta, ainda assim é preservada a vantagem da *Internet* de acessar a um custo mais baixo e de forma mais rápida um número maior de potenciais respondentes.

Valendo-se do fato de que tende a ser relativamente barato contatar muitas pessoas pela *Internet*, Chipp, Goldman e Kleyn (2007) sugerem que se empregue a técnica de saturação (proposta por Turner, 1989, para pesquisas tradicionais), ou seja, de se selecionarem para participar da pesquisa todos os membros do quadro amostral, tornando, assim, irrelevante a discussão sobre amostragem probabilística *versus* não probabilística. Como argumentado pelos autores, o foco deixa de ser a

adequada seleção dos respondentes, baseada no quadro amostral e passa a ser o aprimoramento da definição do próprio quadro amostral.

A saturação, contudo, não resolve o problema do grau de representatividade dos internautas em relação à população-alvo, em geral, (nos casos em que há diferenças entre população e quadro amostral por alguns membros daquela não terem acesso à *Internet*), mas apenas evita os erros adicionais de a amostra selecionada poder ter um perfil de respostas diferente do dos membros do quadro amostral.

A possibilidade de generalização para a população em geral – incluindo-se aqueles membros desta que não têm acesso à *Internet* – depende da verificação sobre se o padrão de respostas de não usuários seria similar a (ou não estatisticamente diferente de) aquele de usuários da *Internet* e também do grau de representatividade da amostra selecionada em relação à totalidade de internautas (tanto acessíveis quanto não acessíveis), que comporiam a população-alvo da pesquisa (BEST; KRUEGER; HUBBARD; SMITH, 2001).

Vale mencionar que a composição gradual de amostras, em que novos elementos vão sendo adicionados até que a distribuição de determinadas variáveis (por ex. demográficas) da amostra de internautas colhida se assemelhe à distribuição já conhecida de amostras coletadas probabilisticamente por outros meios (por ex. telefone), assumindo assim que a amostra poderia ser considerada representativa da população em geral, que vem sendo utilizada por muitos autores (BEST et al., 2001) como forma de gerar amostras “representativas”, pode ser criticada. Além das variáveis utilizadas para comparação entre as amostras, pode haver outras que estariam diferentemente associadas a internautas *versus* não internautas e que poderiam também estar associadas às variáveis de interesse da pesquisa, o que violaria a possibilidade de generalização imediata para a população.

Schillewaert, Langerak e Duhamel (1998) avaliaram o impacto de quatro diferentes métodos de amostragem não probabilística para pesquisas *online*: anúncio de uma página em um periódico semanal de negócios, chamada colocada em uma comunidade da rede, *hyperlinks* apresentados em *websites* que não o da pesquisa, e *e-mail*, não encontrando nenhum efeito estatisticamente significativo sobre a taxa de resposta. Entretanto, os autores identificaram importantes diferenças na composição das amostras em termos do padrão de distribuição das variáveis demográficas ou webográficas (como se o tipo de canal de recrutamento atraísse um determinado perfil específico de usuário), embora não tenham encontrado diferenças em termos das variáveis

atitudinais (relativas ao uso da *Internet*) dos componentes de cada tipo de amostra.

Erros de não resposta em pesquisas online

Uma das principais causas de frustração de respondentes, ao participarem de pesquisa via *webpage*, reside numa apresentação ruim do questionário *online* (SUE; RITTER, 2007), o que pode levar ao aumento das taxas de não resposta.

Dillman e Bowker (2001) sugerem que dificuldades para o usuário responder à pesquisa podem levar à desistência. Os autores reportaram terem observado diversos tipos de frustração por parte dos respondentes, entre os quais: falta de experiência com computadores e com *Internet*, o que dificulta o entendimento do uso de *radio buttons* (que exigem que se clique em outra resposta, caso se queira trocar de resposta, mas que não permitem simplesmente apagar-se todas as respostas após uma ter sido selecionada) e de menus *drop-down* (que apresentam apenas uma caixa com uma breve orientação para o usuário (por ex. “escolha uma opção”), devendo o usuário clicar no lado direito da caixa para abrir as diversas opções disponíveis para resposta, das quais deve ser escolhida apenas uma); dificuldade para enxergar todas as opções de resposta sem rolar a tela; obrigação de responder a todas as perguntas, mesmo quando alguma delas não apresente opções de resposta que pareçam apropriadas; desconhecimento de quanto ainda falta para completar o questionário; obrigação de tomar múltiplas ações para se responder a uma pergunta (por ex. clicar numa opção de resposta, deslocar-se para a barra de rolamento para se poder enxergar um botão de “próxima página”, o qual deve ser clicado para o acesso às perguntas na página seguinte).

O erro de não resposta ocorre porque alguns dos membros selecionados para compor a amostra podem não estar dispostos ou podem não ser capazes de responder à pesquisa. No caso de pesquisas na *Internet*, nas quais não se defina claramente o quadro amostral (por ex. em pesquisas acessadas por meio de um convite aberto em um portal) ou a probabilidade de seleção, não é possível estimar o viés de não resposta (COUPER, 2000). No entanto, se forem solicitadas no início da pesquisa informações “demográficas”, que permitam classificar os participantes, seria possível, ao menos, verificar se haveria diferenças estatisticamente significativas nestas variáveis entre aqueles que responderam a pesquisa até o fim *versus* aqueles que desistiram no meio (*dropouts*), mas não *versus* aqueles que nem iniciaram a resposta (COUPER, 2000). Por outro lado, Dillman (2006) recomenda que tais tipos de perguntas sejam realizados somente ao final da pesquisa, uma vez que elas tendem a não

despertar o interesse dos potenciais respondentes e, em consequência, podem levar ao aumento da desistência se forem colocadas no início – este artifício impediria, no entanto, a comparação entre respondentes e desistentes.

Como recurso para diminuir a taxa de desistência, Dillman, Tortora e Bowker (1998) sugerem que se permita ao respondente prosseguir com as perguntas, mesmo que ele não tenha respondido a todas as perguntas antecedentes, mas, naturalmente, o respondente deve ser alertado, ao tentar submeter o questionário, sobre a eventual existência de perguntas ainda não respondidas. Como mecanismo adicional para manter a disposição para continuar respondendo, os autores sugerem que seja apresentado ao respondente o seu estágio proporcional de completude do questionário, algo que já é atributo implícito nos questionários em papel, mas que também é uma preocupação em outros métodos de pesquisa, tais como entrevistas pessoais ou por telefone.

Erros de mensuração em pesquisas online

Erros de mensuração podem resultar, entre outras causas, do uso de *designs* de questionários *online* que confundam os respondentes, que os induzam a marcar determinadas opções de resposta, ou que tomem muito tempo para serem carregados na tela – o que aumentaria a fadiga e diminuiria a disposição para continuar respondendo. Portanto, questionários simples aumentariam a probabilidade de que todos os “acessados” pela pesquisa tivessem igual chance de responder e oferecessem respostas mais “corretas”. Recomenda-se, portanto, o uso de *layouts* fáceis de serem interpretados e carregados na tela e que produzam o mesmo visual independentemente do equipamento e do *browser* do respondente ou de seus recursos de telecomunicações (DILLMAN; TORTORA; BOWKER, 1998).

Galesic e Bosnjak (2009) verificaram que perguntas localizadas mais para o final do questionário eram respondidas de forma mais rápida, mais curta e mais uniforme (isto é, respostas semelhantes para perguntas diferentes) do que perguntas colocadas no início, possivelmente afetando, portanto, a validade das respostas. Além disso, perguntas mais ao final exibiam maior taxa de respostas em branco.

É preciso levar em conta também que a maioria dos respondentes está acostumada a certa lógica de resposta presente nos questionários em papel. Portanto, algumas particularidades da *Internet* tais como clicar no botão de “enviar” após ter terminado de responder, podem não ser intuitivas para algumas pessoas, o que aumentaria a taxa de não resposta. Por esta razão, Dillman, Tortora e Bowker (1998) recomendam que o

design leve em conta tanto a lógica de operação de computadores quanto a lógica de como as pessoas esperam que os questionários “funcionem” e que, na medida do possível, as perguntas e respostas sejam apresentadas em formato similar aos dos tradicionais questionários em papel e que instruções específicas sobre as ações esperadas sejam clara e explicitamente apresentadas aos respondentes e devem estar próximas à parte da tela onde a ação deve ser tomada.

Um tipo particular de pergunta, aquela em que é solicitado ao respondente que marque todas as respostas que se aplicam com base em uma lista pré-definida de opções, pode levar a um problema de *satisficing* (KROSNICK, 1991), que resulta de o respondente ler e marcar até um determinado número de opções que ele julga que já “seriam suficientes para as necessidades do pesquisador”, tendo como consequência que as opções mais acima na lista teriam maior probabilidade de serem marcadas que aquelas mais abaixo. Tal problema pode ser parcialmente contornado por meio de um rodízio automatizado da posição das opções de resposta. O *satisficing* não é um problema específico somente da *Internet*, mas pode ser agravado em especial se o número de opções (estejam expostas numa lista de *radio buttons* ou em um menu *drop-down*) não aparecerem de forma completa na mesma tela ou exigirem uma ação de *scroll* (rolamento) da tela. Isto seria mais crítico, ainda, quando se tratasse da primeira pergunta, a qual deveria estar sempre completamente visível na primeira tela (DILLMAN; TORTORA; BOWKER, 1998). Mais ainda, se a mesma pesquisa se utilizar de diferentes métodos de coleta de dados, uma pergunta deste tipo pode não ser facilmente replicável num formato de entrevista. Por esta razão, Dillman, Tortora e Bowker (1998) recomendam que os questionários na *Internet* levem em conta a possibilidade de uso das perguntas em múltiplos métodos de coleta de dados.

Uma das causas de ocorrência de erros de mensuração é que, em função de diferenças em equipamentos, navegadores e conexões de acesso, é possível que diferentes respondentes vejam o *layout* do questionário de forma diferente, sendo que muitos poderão ver um formato ou uma disposição visual distintos daqueles imaginados pelo pesquisador (DILLMAN; BOWKER, 2001). Por ex. as distâncias entre pontos de uma escala de respostas podem mudar em função da resolução da tela ou parecerem diferentes em função da cor de fundo; e a quantidade de perguntas e de opções de resposta que são mostradas em cada tela também variam conforme a resolução da tela no computador do usuário. Esta situação ficou ainda mais agravada com a rápida penetração de novos dispositivos de acesso à *Internet*, tais como telefones de terceira e

quarta geração, como os *iPhones*, e *tablets*, como os *iPads*, com telas menores e com sistemas operacionais específicos.

Quanto ao erro de mensuração decorrente de respostas indevidas, este pode por vezes ser reduzido se for possível incluir verificação das respostas, de forma a informar ao respondente caso estas estejam fora dos valores admissíveis para a respectiva variável. Existe outro tipo de erro de mensuração em pesquisas pela *Internet* quando se utilizam painéis ou pesquisas longitudinais (COUPER, 2000), que é o fato de que, por sua experiência com o instrumento e as questões da pesquisa, os participantes do painel possam ter um padrão de respostas consistentemente distinto daquele de novos (ou menos experientes) participantes (tal problema também ocorre em pesquisas longitudinais tradicionais, mas pode ser agravado na *Internet* pelo fato de que a mudança no endereço eletrônico e, em consequência na composição do painel, é mais frequente do que a do endereço físico).

Dillman e Bowker (2001) alertam para o fato de que, se os respondentes forem obrigados a responder a todas as perguntas anteriores antes de poderem prosseguir para as seguintes, estes podem se sentir desestimulados e virem a desistir – aumentando o viés de não resposta – ou simplesmente proverem uma resposta qualquer – aumentando o erro de mensuração – apenas para conquistarem o direito de seguir adiante.

Uma forma de se identificar um possível tipo de respondente que forneceria respostas “incorretas” é o registro automático, na base de dados, de informação sobre do horário de início e término do questionário. Esta informação permitiria identificar os questionários que foram preenchidos em tempo muito inferior ao padrão da amostra, o que poderia indicar a falta de comprometimento dos entrevistados no preenchimento das respostas, cabendo ao pesquisador decisão de expurgá-los da amostra final. O estudo de Malhotra (2008) indicou que aqueles respondentes que terminaram de responder em um tempo consideravelmente abaixo do dos demais (mais de 1.5 desvios padrão abaixo da média) apresentavam um maior efeito de precedência (*primacy effect*, ou seja, a tendência a selecionar opções de resposta que apareçam primeiro) ou de *satisficing*. Malhotra (2008), contudo, adverte que o pesquisador não deve considerar apressadamente remover estes respondentes “rápidos”, posto que a magnitude do efeito de precedência que ele observou em seu estudo era moderada por outras variáveis, tais como habilidades cognitivas do respondente e tipo de escala de resposta (ordinal vs. categórica). Além disso, pode-se usar a randomização das opções de resposta entre respondentes para minimizar este tipo de efeito.

Dillman e Bowker (2001) resumem uma série de princípios que deveriam ser seguidos em pesquisas *online* bem como os tipos de erros usuais em *surveys* que tais princípios procurariam minimizar (vide Quadro 1).

Quadro 1: Princípios de projeto de pesquisas *online* e sua relação com fontes de erros em *surveys*

	Tipo de erro			
	Erro de cobertura	Erro de amostragem	Erro de não resposta	Erro de mensuração
1. Apresentar o questionário com uma tela de boas-vindas que motive e dê instruções aos respondentes sobre como proceder			X	
2. Prover um código de identificação para limitar o acesso apenas a membros da amostra selecionada	X	X		
3. Escolher como primeira pergunta uma que desperte o interesse dos respondentes e seja facilmente respondida, colocando-a bem visível na primeira tela do questionário			X	
4. Apresentar todas as questões em um formato que seja similar ao que costuma ser usado em pesquisas em papel			X	X
5. Restringir o uso de cores de forma a preservar a legibilidade, o fluxo navegacional e as propriedades de mensuração das opções de resposta				X
6. Evitar diferenças na aparência visual resultante de diferentes configurações de tela, sistemas operacionais, navegadores, telas parciais e quebra de texto	X		X	X
7. Oferecer, no exato ponto em que forem necessárias, instruções claras sobre como executar cada ação no computador para responder às perguntas			X	
8. Minimizar o uso de menus <i>drop-down</i> e identificá-los claramente com uma instrução de “clique aqui para selecionar sua resposta”				X

9. Não exigir dos respondentes que todas as perguntas anteriores tenham sido respondidas antes de se poder prosseguir para as seguintes		X	X
10. Oferecer orientações sobre eventual “pulo” de questões que não se apliquem			X
11. Construir o questionário <i>online</i> de forma que seja possível pagnar de uma pergunta para outra, a menos que uma estrita exigência de ordem seja necessária ou que sejam paralelamente utilizados outros métodos de coleta (por ex. telefone)	X	X	X
12. Nos casos em que o número de opções de resposta exceda o tamanho de uma tela, se apresentadas em uma única coluna, considerar o uso de duas colunas ou outra forma apropriada de agrupamento e encadeamento entre pergunta e opções de resposta			X
13. Utilizar símbolos gráficos ou frases que explicitem o ponto em que o usuário se encontra no progresso em direção à conclusão, evitando, contudo, os recursos que tomem muito tempo para carregamento	X	X	
14. Utilizar símbolos gráficos ou frases que explicitem o ponto em que o usuário se encontra no progresso em direção à conclusão, evitando, contudo, os recursos que tomem muito tempo para carregamento	X	X	
15. Evitar formatos de questões sabidamente problemáticas em pesquisas tradicionais, tais como “marque todas as opções que se aplicam” ou perguntas com respostas abertas		X	X

Fonte: Dillman e Bowker (2001).

As orientações #8 e #10 referem-se a questões específicas de pesquisas na *Internet*, enquanto os aspectos relacionados a erros de mensuração abordados nas orientações #4, #12 e #14 se referem a

potenciais problemas de compatibilidade quando a pesquisa se vale de múltiplos métodos de coleta de dados.

IMPACTO DO FORMATO DO QUESTIONÁRIO SOBRE A QUALIDADE DAS RESPOSTAS

O formato do questionário, seu *design*, bem como a organização e sequenciamento das perguntas e das opções de resposta podem levar a erros sistemáticos de mensuração, os chamados vieses de método. Embora o viés de método (PODSAKOFF; MACKENZIE; LEE, 2003) possa ocorrer também em pesquisas tradicionais, algumas decisões sobre o formato do questionário, que sejam específicas de pesquisas *online*, podem ter influência sobre o padrão de respostas, em especial pelo fato de que, na ausência de um entrevistador para motivar e esclarecer dúvidas, o respondente busca auxílio nas próprias “pistas” providas pelo instrumento, tais como seus elementos verbais e visuais (SCHWARZ, 1995).

Entre os diversos aspectos relacionados ao formato e *design* de questionários *online*, os seguintes serão aqui discutidos: a) uso de respostas forçadas; b) vantagens e desvantagens de três tipos específicos de formatos para registro das respostas, especificamente *radio buttons*, menus *drop-down* e caixas de entrada (*entry boxes*); c) a influência do uso de palavras, símbolos ou itens gráficos em *web surveys*; d) tamanho do questionário e apresentação de indicador de progresso; e e) proximidade e agrupamento das questões.

Uso de respostas forçadas

Respostas “forçadas” obrigam o usuário a selecionar uma das opções de resposta oferecidas no questionário de forma a permitir que se avance para a próxima pergunta (ou a responder a todas as perguntas de uma dada página antes de prosseguir para a página seguinte). Esta exigência tem por objetivo evitar dados ausentes, mas pode acarretar a desistência de alguns respondentes – com o correspondente possível aumento do viés de não resposta – ou forçar outros a darem uma resposta que eles não julgariam como a mais apropriada – aumentando, assim, o erro de mensuração. Uma opção para minimizar esta fonte de erro de mensuração é oferecer uma opção de resposta do tipo “não sei” ou “não se aplica” ou “prefiro não responder”. Assim, garante-se que a resposta à pergunta não teria sido deixada em branco por esquecimento, mas não se força uma resposta incorreta. Albaum et al. (2010) investigaram o impacto sobre a taxa de completude (*completion rate*) dos questionários e a taxa de

respostas em branco decorrentes do uso, isolado ou em conjunto, de respostas forçadas ou da opção “prefiro não responder”. Seus resultados sugerem que o emprego de respostas forçadas parece não ter relação com a taxa de completude dos questionários *online*. Similarmente, não foi encontrada evidência de que o uso da opção “prefiro não responder” afetaria a taxa de completude, independentemente de o questionário conter ou não algumas perguntas com resposta forçada. Nesse estudo, em particular, a opção “prefiro não responder” foi utilizada em volume suficiente a ponto de indicar que, no caso de serem usadas respostas forçadas, os usuários teriam registrado, em várias perguntas, uma opção de resposta que não corresponderia à realidade. Embora a escolha (forçada) de uma opção de resposta diminua a taxa de resposta em branco, por outro lado, de nada adianta a presença de uma resposta se esta não é válida, levando, assim, a erro de mensuração. Conforme argumentado por Albaum et al. (2010), o uso de opções de resposta do tipo “não sei”, “não se aplica” ou “prefiro não responder” representa um *trade-off* entre quantidade *versus* qualidade dos dados.

O estudo de Heerwegh e Loosveldt (2002) indicou que a oferta de uma opção de resposta em branco (que eles esperavam que reduzisse a frustração dos respondentes, levando a menores taxas de desistência) não mostrou efeito estatisticamente significativo sobre esta.

Uso de radio buttons, menus drop-down e caixas para digitação de texto (entry boxes)

Radio buttons se apresentam em um formato semelhante ao que as pessoas estão acostumadas a ver em perguntas com opções fechadas de respostas em questionários tradicionais e só permitem que se escolha uma única opção de resposta; contudo, uma vez marcada uma opção de resposta, esta pode ser trocada por outra, mas não é mais possível apagar todas as opções e fornecer uma resposta em branco, a menos que o pesquisador tenha incluído explicitamente uma opção de resposta “em branco” (ex: “não sei” ou “não se aplica”). Vale ressaltar que a opção de resposta em branco deve ser permitida sempre que houver a possibilidade de os respondentes não possuírem ou não desejarem fornecer a informação solicitada na pergunta.

Menus *drop-down*, por seu turno, apresentam apenas uma caixa com uma breve orientação (por ex. “escolha uma opção”), devendo o usuário clicar no lado direito da caixa para abrir as diversas opções disponíveis para resposta; pode ser selecionada apenas uma única opção, mas, em geral, o usuário pode optar por uma resposta em branco, mesmo após ter selecionado previamente outra opção (basta que ele clique novamente na

frase com a orientação geral, desde que esta não tenha sido, indevidamente, definida como uma das opções de resposta; neste caso, esta se tornaria a opção *default* de resposta, o que poderia enviesar a resposta e levar a erro de mensuração).

Já *entry boxes* ou caixas de entrada são espaços na tela para que usuário digite em texto livre sua resposta.

Heerwegh e Loosveldt (2002) notaram que, uma vez que menus *drop-down* ocupam menos espaço na tela do que opções (se verticalmente alinhadas) de resposta em formato *radio button*, o primeiro formato permitiria apresentar mais perguntas numa mesma tela. Por seu turno, Dillman (2006) observou que menus *drop-down* podem desorientar os respondentes, em especial quando se escolhe uma opção de resposta que esteja na parte inferior da lista, uma vez que, ao se fechar automaticamente o menu *drop-down*, o ponteiro do *mouse* estará posicionado várias linhas abaixo da pergunta respectiva, podendo o respondente, inadvertidamente, pular as perguntas localizadas entre aquela recém-respondida e a atual localização do ponteiro do *mouse*.

Heerwegh e Loosveldt (2002) verificaram que os usuários levavam mais tempo para responder se o formato era um menu *drop-down*, o qual também se mostrou associado a uma maior taxa de desistência. Contudo, não houve diferença estatisticamente significativa na taxa de respostas em branco associadas com menus *drop down* ou *radio buttons*. Com base em seus resultados, os autores sugerem que se utilizem *radio buttons*, em especial no caso de usuários menos proficientes com o meio *online*.

Couper, Traugott e Lamias (2001) verificaram se o uso de *radio buttons* ou *entry boxes* afetaria o tempo para completar a pesquisa, a taxa de dados ausentes e taxa de respostas inválidas. Os autores testaram também o efeito do comprimento da caixa de entrada. Nos questionários com caixa de entrada, os autores davam a opção de os respondentes deixarem a caixa de entrada em branco ou marcarem a respectiva opção de “não sei”, enquanto que nos questionários com *radio buttons* os respondentes teriam que selecionar obrigatoriamente uma opção (ainda que fosse a “não sei”). Os autores não observaram diferença estatisticamente significativa no tempo total para preenchimento dos questionários. Por seu turno, a taxa de dados ausentes foi estatisticamente maior quando do uso de caixa de entrada, mas não se observaram diferenças associadas ao tamanho ou à forma da caixa de entrada. Quanto às respostas inválidas, por definição, elas não são possíveis quando do uso de *radio buttons*, mas o estudo mostrou que uma caixa de entrada mais longa tendia a levar os respondentes a digitarem respostas inválidas (ou

seja, não aceitáveis, dado o tipo de pergunta e as possibilidades esperadas de resposta) que uma caixa de entrada mais curta, sugerindo que o tamanho da caixa para digitação influencia a quantidade de informação que os respondentes julgam que o pesquisador espera.

Smyth, Dillman, Christian e McBride (2009) investigaram se a qualidade das respostas, fornecidas a perguntas abertas, dependeria das características da caixa de entrada destinada à resposta – especificamente, o tamanho da caixa ou a inclusão ou não de explicações de que a respostas poderiam exceder o tamanho original da caixa – ou da apresentação de uma introdução explicativa e motivadora na pergunta. Seus resultados indicam que o tamanho da caixa de resposta parece não influenciar a qualidade das respostas dos respondentes iniciais (*early respondents*), mas aumentaria significativa a qualidade das respostas fornecidas pelos respondentes tardios (*late respondents*), que seriam, em princípio, aqueles menos motivados a responder. Contudo, a instrução de que a extensão da resposta não precisaria se limitar ao tamanho original da caixa, bem como a inclusão de uma introdução que enfatizasse a importância da qualidade e do tamanho da resposta se mostraram associadas ao aumento da qualidade das respostas (em termos de extensão das mesmas, quantidade de temas abordados, tempo dedicado à resposta e redução de respostas em branco), tanto de respondentes iniciais quanto de respondentes tardios.

Influência do uso de palavras, símbolos ou itens gráficos em web surveys

Assim como em pesquisas tradicionais em papel, também nas pesquisas *online* a qualidade das respostas pode ser afetada pelo uso de números, símbolos e sinais gráficos em complemento às palavras (CHRISTIAN; DILLMAN, 2004; TOURANGEAU; COUPER; CONRAD, 2004).

Christian, Dillman e Smyth (2007) investigaram se diferentes aspectos visuais do questionário – em particular, tamanho dos espaços destinados às opções de resposta, uso de símbolos ao invés de palavras, tipo de linguagem verbal empregada na pergunta e localização da instrução aos respondentes – estariam associados ao aumento da probabilidade de que os respondentes fornecessem respostas em um formato que seria desejado pelo pesquisador (por ex. dois dígitos para o mês e quatro dígitos para o ano), evitando-se a necessidade de emissão de mensagens de erro ou de solicitação de correção aos respondentes. Dessa forma, poderia ser reduzida a frustração e possível abandono do questionário por parte dos respondentes (BEST; KRUEGER, 2004). O estudo de Christian, Dillman e Smyth (2007) indicou que os respondentes apresentavam maior probabilidade de registrar a informação no formato

desejado pelo pesquisador (neste caso, dois dígitos para o mês e quatro dígitos para o ano) nas seguintes condições: espaço menor para o mês e maior para o ano, uso de símbolos (MM e YYYY) ao invés de palavras (“Month” e “Year”), maior proximidade de símbolo em relação ao espaço para preenchimento da respectiva informação. Por outro lado, a localização específica do símbolo (por ex. à esquerda, à direita, abaixo ou acima da caixa para resposta) não pareceu exercer impacto significativo sobre a probabilidade de o respondente usar o formato desejado e a forma e detalhamento da linguagem verbal empregada na pergunta não apresentaram efeito significativo ao menos naqueles casos em que o questionário já apresentava instruções apropriadas de preenchimento (tal como discutido anteriormente) junto aos espaços de resposta. Os resultados do estudo sugerem que, além dos cuidados usuais com a linguagem da pergunta, os pesquisadores que se valem de *web surveys* devem também ficar atentos para a apresentação visual e as orientações relativas às opções de resposta. Christian, Dillman e Smyth (2007) ponderaram também sobre a possibilidade de menus *drop down* ou de calendários – em que o respondente clicaria sobre a opção de resposta – serem inapropriados, posto que alguns respondentes podem não ter familiaridade com este tipo de interface gráfica, além do fato de que este formato apresenta um número grande de opções para escolha (por ex. de 1900 a 2000 para o ano).

Tamanho do questionário e apresentação de indicador de progresso

Meta-análises sobre o efeito do tamanho do questionário em papel na propensão a responder indicam que, quanto maior o tamanho, em geral, menor a taxa de resposta (YAMMARINO; SKINNER; CHILDERS, 1991).

Contudo, ainda há poucos estudos que avaliaram este efeito em *web surveys* (GALESIC e BOSNJAK, 2009). O impacto do tamanho pode depender de como os participantes são informados a respeito: se previamente ao início das respostas ou se “descobrem esta informação” diretamente somente à medida que vão avançando no questionário. Embora a meta-análise de Cook, Heath e Thompson (2000) tenha indicado que o tamanho do questionário não afetaria a taxa de resposta, alguns estudos indicaram que quanto maior o tempo para preenchimento anunciado *a priori* (CRAWFORD; COUPER; LAMIAS, 2001; GALESIC; BOSNJAK, 2009; MARCUS et al., 2007) ou inferido / observado pelos respondentes (DEUTSKENS et al., 2004), menor a taxa de resposta (ou seja, a percentagem daqueles que efetivamente iniciam as respostas). Galesic e Bosnjak (2009) observaram também que, quanto

menor o tamanho, menor a taxa de abandono daqueles que iniciaram o processo de respostas.

Couper, Traugott e Lamias (2001) investigaram se a presença de um indicador de progresso teria impacto sobre a taxa de desistência, o tempo médio para completar o questionário e a taxa de dados ausentes (que, em conjunto com a taxa de completude, está relacionado ao viés de nãoresposta). Ao contrário do esperado pelos autores, a inclusão de um indicador de progresso não mostrou ter impacto estatisticamente significativo sobre o percentual de questionários completados. Já o tempo médio para completar a pesquisa foi maior no caso de questionários com o indicador de progresso, o que foi atribuído pelos autores ao tempo adicional necessário para carregar os elementos gráficos associados a tal indicador (naturalmente, à medida que mais usuários vão tendo acesso à banda larga, tal tempo tende a ter seu impacto reduzido). O indicador de progresso também não se mostrou associado à presença de dados ausentes nessa pesquisa específica, em que sempre era exigido que os participantes escolhessem uma opção de resposta para cada pergunta, ainda que fosse marcando a opção “não sei”, o que significa que eles teriam o mesmo “trabalho” quer marcassem esta opção quer marcassem outra.

Proximidade e agrupamento das questões

Com relação ao agrupamento ou não em uma mesma tela de todas as questões relacionadas entre si (por ex. indicadores de um mesmo construto), Couper, Traugott e Lamias (2001) investigaram se haveria impacto sobre as correlações observadas, o tempo para completar o questionário e a taxa de dados ausentes. Como esperado por eles, as correlações entre os itens (medidos pelo alfa de Cronbach) foram maiores quando do agrupamento em uma mesma tela, o que sugere que o arranjo físico do questionário afeta o padrão de respostas. O tempo total médio para completar o questionário foi estatisticamente menor quando do agrupamento, embora não fosse possível observar diferenças estatisticamente significativas nos tempos de resposta a cada um dos subconjuntos de itens agrupados em uma mesma tela *versus* desagregados em telas sequenciais. Já o agrupamento se mostrou associado a uma taxa estatisticamente menor de dados ausentes (respostas do tipo “não sei” ou “não se aplica”). Os autores haviam especulado, contudo, que poderiam ser esperados dois efeitos de impacto contrário do agrupamento: se, por um lado, o agrupamento tenderia a reduzir o esforço de tempo, o que tenderia a aumentar o índice de respostas substantivas (em detrimento de “não sei”), por outro lado o agrupamento poderia induzir ao uso de um mesmo padrão de respostas para todas as perguntas em uma mesma tela

de tal sorte que, se fosse escolhida uma resposta “não sei”, poderia haver uma tendência a escolher outras respostas “não sei” para as demais perguntas na mesma tela. Os autores alertam, ainda, que o efetivo agrupamento de perguntas em uma mesma tela depende da configuração do computador do respondente e pode ser-lhe necessário rolar a tela mesmo que o pesquisador tenha projetado as perguntas para aparecerem juntas na mesma tela.

QUESTÕES ÉTICAS EM PESQUISAS COM O USO DA *INTERNET*

Não é possível garantir total anonimato em pesquisas por meio da *Internet*. Se a resposta for encaminhada por meio de um *e-mail*, o respondente estará sendo implicitamente identificado. Mesmo que o questionário seja preenchido em um *website*, ainda assim é possível ter informação ao menos sobre o computador do respondente, por meio de seu IP (*Internet Protocol*). Claramente, alguns dos cuidados sugeridos para se evitarem múltiplas respostas por um mesmo respondente (tais como senha ou identificação embutida no *link*) podem tornar a identificação do respondente ainda mais transparente. Esta situação pode levar ao aumento da taxa de não resposta, se os potenciais respondentes se sentirem desconfortáveis ou intimidados, ou a aumento do erro de mensuração, em especial em situações que envolvam viés de conveniência social (*social desirability bias*), isto é, o registro de uma resposta que seria considerada “apropriada” ou “socialmente correta”, em lugar da resposta verdadeira.

Torna-se, portanto, relevante informar aos potenciais respondentes que medidas foram tomadas para evitar o acesso inadvertido de terceiros aos dados da pesquisa (por ex. com o uso de conexões SSL – *secure sockets layer*) (STANTON; ROGELBERG, 2001), que as respostas individuais serão mantidas confidenciais e que serão reportados somente resultados agregados. Contudo, não deve ser feita qualquer promessa de total anonimato (SIMSEK; VEIGA, 2001), uma vez que tal não pode ser garantido em pesquisas *online*.

Por questões éticas, sempre que uma pesquisa for deixar gravado um *cookie* no computador do usuário, tal fato deveria ser deixado claro ao usuário e deveria ser-lhe explicado por que a gravação do *cookie* seria necessária como forma de garantir a “confiabilidade” da pesquisa (SIMSEK; VEIGA, 2001).

CONSIDERAÇÕES FINAIS

A *Internet*, como meio de coleta de dados, oferece uma série de recursos que podem ser úteis em *surveys*, ao mesmo tempo em que também traz desafios que lhe são peculiares (SCHONLAU; FRICKER JR.; ELLIOT, 2002). Pode haver relevantes diferenças entre membros da população-alvo, tanto no efetivo acesso à *Internet* quanto em sua proficiência com o uso de recursos desse meio, como também diferenças na infraestrutura (*hardware* e *software*) dos respondentes (BRADLEY, 1999), o que pode levar a erros de cobertura, de não resposta e de mensuração.

Decisões relativas à definição do quadro amostral, às técnicas de amostragem, ao formato dos questionários e ao tratamento de questões éticas podem ter importantes implicações sobre a qualidade dos dados coletados, em particular sobre o tamanho da amostra, seu grau de representatividade, o erro de mensuração, a taxa de resposta, a presença de dados ausentes ou o padrão de distribuição das respostas. Entre aquelas pessoas que possuem acesso à *Internet*, é necessário encontrar um mecanismo que garanta a probabilidade conhecida de se selecionar cada respondente potencial (DILLMAN; TORTORA; BOWKER, 1998).

Conforme alertado por Dillman e Bowker (2001), a simples obtenção de uma amostra grande não é por si só garantia de “qualidade” dos dados, em especial se os respondentes tiverem caráter voluntário. Os autores complementam afirmando que ignorar a necessidade de se definir a população de interesse à pesquisa, de se selecionar uma amostra aleatória e de se obter altas taxas de respostas (sobre a amostra selecionada) representa uma grave ameaça à validade de pesquisas *online*.

Alguns dos problemas de pesquisas com o uso da *Internet*, como os relacionados a erros de cobertura por falta de acesso ou por ausência de uma lista de todos os membros da população-alvo, podem ser minimizados quando tais pesquisas ocorrem dentro de uma dada organização. Contudo, o viés de não resposta continua merecendo cuidados conforme apontado neste estudo.

A habilidade de cativar o potencial respondente de forma a convencê-lo a participar da pesquisa, que é um aspecto importante em diversos métodos de coleta de dados, torna-se crítico em pesquisas *online*, uma vez que o internauta tem sua atenção atraída por muitos outros estímulos e pode muito facilmente desviar seu interesse para outro *site* ou tarefa em seu computador (SIMSEK; VEIGA, 2001). Portanto, a forma como o pesquisador se apresenta, bem como o tema de sua pesquisa e a entidade patrocinadora é de extrema importância para aumentar a taxa de

resposta.

As particularidades da *Internet* e as dúvidas quanto à sua adequabilidade (ainda manifestadas por alguns acadêmicos) podem exigir do pesquisador um esforço adicional para demonstrar o rigor de seu método de coleta de dados, que pode ser conseguido, ainda que a um custo extra, por meio do emprego de validação cruzada com uma amostra tradicional (STANTON; ROGELBERG, 2001).

Vale dizer que, para a elaboração do questionário *online*, devem ser observadas algumas das mesmas recomendações da sua contraparte em papel: deve ter aparência profissional e estimulante, ser de fácil compreensão e não intimidar os respondentes, tornar claro e fácil o processo de responder às questões e ser acessível a todos na população-alvo ou no quadro amostral considerado adequado pelo pesquisador.

O instrumento *online* pode ser até mesmo o mais apropriado quando a pesquisa tiver algumas características: (1) se a amostra for relativamente grande e estiver bastante distribuída geograficamente; (2) se existir a necessidade de resultados rápidos; (3) quando a questão do anonimato não for tão relevante, podendo ser melhor tratada via *webpage* do que por *e-mail*; e (4) quando os respondentes tiverem acesso à *Internet* e, nestes casos, a taxa de respostas puder tornar-se maior quando os respondentes forem contatados antes por *e-mail* (SUE; RITTER, 2007).

Afinal, como Dillman (2006) afirmou, tanto o correio quanto o telefone foram considerados inicialmente pobres alternativas para a entrevista pessoal, sendo o mesmo dito mais recentemente sobre as ferramentas de pesquisa via *Internet*. Acredita-se que é preciso não assumir a *Internet* como solução geral para *surveys*, mas sim tomar os devidos cuidados para avaliar a aplicabilidade do meio *online* e suas particularidades, vantagens e limitações. Espera-se que a discussão aqui apresentada sobre particularidades das pesquisas via *webpage* e os problemas enfrentados e as soluções adotadas nos estudos analisados sirvam de orientação e incentivo para que pesquisadores brasileiros se sintam mais confiantes quanto ao uso apropriado da coleta de dados *online* em suas próprias pesquisas.

REFERÊNCIAS

ALBAUM, G.; ROSTER, C.; WILEY, J.; ROSSITER, J.; SMITH, S. Designing Web Surveys in Marketing Research: Does Use of Forced Answering Affect Completion Rates? *Journal of Marketing Theory and Practice*, v.18, n.3, p.285–293, 2010.

BEST, S.; KRUEGER, B.; *Internet data collection (Quantitative Applications in the Social Sciences)*. Thousand Oaks, CA: Sage, 2004.

BEST, S.; KRUEGER, B.; HUBBARD, C.; SMITH, A. An assessment of the generalizability of Internet surveys. *Social Science Computer Review*, v.19, n.2, p.131-145, 2001.

BRADLEY, N. Sampling for Internet surveys: an examination of respondent selection for Internet research. *Journal of the Market Research Society*, v.41, n.4, p.387-395, 1999.

CHIPP, K.; GOLDMAN, M.; KLEYN, N. What they really think: resolving methodological issues in supply chain ethics research. *Acta Commercii*, v.7, p.112-122, 2007.

CHRISTIAN, L.; DILLMAN, D. The Influence of Symbolic and Graphical Language Manipulations on Answers to Paper Self-Administered Questionnaires. *Public Opinion Quarterly*, v.68, p.57-80, 2004.

CHRISTIAN, L.; DILLMAN, D.; SMYTH, J. Helping Respondents Get it Right the First Time: The Influence of Words, Symbols, and Graphics in Web Surveys. *Public Opinion Quarterly*, v.71, n.1, p.113-125, 2007.

COOK, C.; HEATH, F.; THOMPSON, R. A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement*, v.60, p.821-36, 2000.

COUPER, M. Web surveys: a review of issues and approaches. *Public Opinion Quarterly*, v.64, n.4, p.464-494, 2000.

COUPER, M.; TRAUGOTT, M.; LAMIAS, M. Web survey and administration. *Public Opinion Quarterly*, v.65, n.2, p.230-253, 2001.

CRAWFORD, S.; COUPER, M.; LAMIAS, M. Web Surveys: Perceptions of Burden. *Social Science Computer Review*, v.19, p.146-62, 2001.

DEUTSKENS, E., De RUYTER, K., WETZELS, M.; OOSTERVELD, P. Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study. *Marketing Letters*, v.15, p.21-36, 2004.

DIB, L.; ROCHA, A.; SILVA, J. The internationalization process of Brazilian software firms and the born global phenomenon: Examining firm, network, and entrepreneur variables. *Journal of International Entrepreneurship*, v. 8, p. 233-253, 2010.

DILLMAN, D. *Mail and Internet surveys: the tailored design method*. 2nd ed., New York: John Wiley & Sons, 2006.

DILLMAN, D.; BOWKER, D. The web questionnaire challenge to survey methodologists. In: REIPS, U.-D.; BOSNJAK, M. (Eds.) *Dimensions of Internet Science*. Lengerich, Germany: Pabst Science Publishers, 2001.

DILLMAN, D.; TORTORA, R.; BOWKER, D. *Principles for constructing web surveys*. Technical Report, 98-50, Social and Economic Sciences Research Center Washington: Pullman, 1998. Disponível em: <http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf>. Acesso em: 20/04/2009.

ENANORIA, W. *Sampling Frames and Coverage Error*. (s/d). Disponível em: www.idready.org/courses/2005/spring/survey_SamplingFrames.pdf. Acesso em: 05/10/2011.

EVANGELISTA, F. Qualitative insights into the international new venture creation process. *Journal of International Entrepreneurship*, v.3, p.179-198. 2005.

GALESIC, M.; BOSNJAK, M. Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, v.73, n.2, p.349-360, 2009.

HAIR, JR., J.; BLACK, W.; BABIN, B.; ANDERSON, R.; TATHAM, R. *Multivariate data analysis*. 6th ed., Upper Saddle River, NJ: Prentice Hall, 2006.

HEERWEGH, D.; LOOSVELDT, G. An evaluation of the effect of response formats on data quality in web surveys. *Social Science Computer Review*, v.20, n.4, p.471- 484, 2002.

KROSNICK, J.. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, v.5, p.213-36, 1991.

MALHOTRA, N. Completion time and response order effects in web surveys. *Public Opinion Quarterly*, v.72, n.5, p.914-934, 2008.

MALHOTRA, N. *Pesquisa de marketing: uma orientação aplicada*. Porto Alegre: Bookman. 2006.

MARCUS, B.; BOSNJAK, M.; LINDNER, S.; PILISCHENKO S.; SCHUETZ, A. Compensating for Low Topic Interest and Long Surveys:

A Field Experiment on Nonresponse in Web Surveys. *Social Science Computer Review*, v.25, p.372–83, 2007.

PODSAKOFF, P.; MacKENZIE, S.; LEE, J. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, v.88, p.879-903, 2003.

SCHILLEWAERT, N.; LANGERAK, F.; DUHAMEL, T. Non-probability sampling for WWW surveys: a comparison of methods. *Journal of the Market Research Society*, v.40, n.4, p.307-322, 1998.

SCHONLAU, M.; FRICKER JR., R.; ELLIOT, M. *Conducting research surveys via e-mail and the web*. Santa Monica, CA, Rand, 2002.

SCHWARZ, N. What respondents learn from questionnaires: the survey interview and the logic of conversation. *International Statistics Review*, v.63, n.2, p.153-168, 1995.

SIMSEK, Z.; VEIGA, J. A primer on Internet organizational surveys. *Organizational Research Methods*, v.4, n.3, p.218-235, 2001.

SMYTH, J.; DILLMAN, D.; CHRISTIAN, L.; MCBRIDE, M. Open-Ended Questions in Web Surveys: Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality? *Public Opinion Quarterly*, v.73, n.2, p.325–337, 2009.

STANTON, J.; ROGELBERG, S. Using Internet/intranet web pages to collect organizational research data. *Organizational Research Methods*, v.4, n.3, p. 200-217, 2001.

SUE, V.; RITTER, L. *Conducting online surveys*. California: Sage Publications. 2007.

TOLEDANO, Y. Use Online Surveys to Get the Feedback You Need: Tools and best practices for conducting Web surveys. 2008. Disponível em:
<http://www.techsoup.org/learningcenter/internet/archives/page9207.cfm>. Acesso em: 05/10/2011.

TOURANGEAU, R.; COUPER, M.; CONRAD, F. Spacing, Position, and Order. Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly*, v.68, p.368–93, 2004.

TURNER, W. Small business data collection by area censuring: a field test of saturation surveying methodology. *Journal of the Market Research Society*, v.31, n.2, p.257-272, 1989.

YAMMARINO, F.; SKINNER, S.; CHILDERS, T. Understanding Mail Survey Response Behavior: A Meta-Analysis. *Public Opinion Quarterly*, v.55, p.613–39, 1991.

APÊNDICE A – FERRAMENTAS PARA CONFEÇÃO DE QUESTIONÁRIOS ONLINE

A título de complemento a este artigo, vale ressaltar que existem diversas ferramentas disponíveis na *internet* para a confecção de questionários *online* e a condução de *web surveys*, tais como: eSurveysPro, Google Docs, kiwksurveys, MakeSurvey, Opinio Survey, QuestionPro, SpeedSurvey, SurveyGizmo, SurveyGold, SurveyMonkey, ZipSurvey, Zoomerang, entre outras.

Os interessados podem consultar uma lista de mais de 300 ferramentas em:

<http://www.websm.org/dbarchive/18/Software/?&p1=1123&p2=82&p3=85&type=0>.

Tais ferramentas costumam ter versões pagas (que oferecem maior variedade de opções e permitem um uso mais intenso da ferramenta), bem como versões gratuitas (as quais, porém, são em geral suficientes para atender a boa parte das necessidades dos pesquisadores).

A maioria destas ferramentas oferece facilidades de apresentação *online* dos resultados, acesso a tutoriais e a FAQ (*frequently asked questions*). Contudo, tais ferramentas podem variar consideravelmente entre si (conforme ilustrado por Toledano, 2008) em termos de:

- Número máximo de questionários que podem ser administrados;
- Número máximo de questões ou de opções de repostas por questionário;
- Prazo em que os resultados ficam disponíveis para consulta pelo criador da *survey*;
- Preços, políticas de desconto (ex: indivíduos vs. empresas), periodicidade de renovação;
- Variedade de diferentes tipos de questões;
- Oferta de questionários padrão *versus* possibilidade de criação de questionários customizados;
- Permissão para inclusão de imagens ou logos;
- Possibilidade de randomização da ordem das perguntas e das opções de resposta;
- Possibilidade de verificação (validação) *online* das respostas;
- Possibilidade de salto lógico entre questões conforme resposta dada a questões anteriores;

-
- Possibilidade de uso de *piping* (propagação), ou seja, carregar a resposta a uma questão ou algum parâmetro da pesquisa para outra questão seguinte ou outra parte do questionário, o que poder tornar a pesquisa mais dinâmica, ágil e personalizada;
 - Recursos para salvar, copiar ou editar questionários anteriormente criados;
 - Possibilidade de personalizar a página de agradecimentos;
 - Suporte a múltiplos idiomas;
 - Ocultação do nome da ferramenta de administração de *web surveys*;
 - Aderência ao padrão 508 dos EUA (<http://www.section508.gov/>), referente a requisitos e normas para facilitar o acesso de pessoas portadoras de deficiência;
 - Possibilidade de uso de listas de *e-mails* do criador da *survey*;
 - Grau de personalização dos convites por *e-mail*;
 - Envio de lembretes aos potenciais respondentes;
 - Agendamento da disponibilização do questionário;
 - Tabulação das respostas;
 - Filtragem e segmentação dos respondentes;
 - Exportação para formato Excel, Word, PowerPoint ou CSV;
 - Geração de apresentações em PowerPoint;
 - Formas de suporte técnico (ex: via *e-mail*, fóruns, *chats* ou telefone).

Os interessados em aspectos metodológicos atinentes a pesquisas com auxílio da *internet* poderão encontrar muitas informações e orientações interessantes em: www.websm.org.

DADOS DOS AUTORES

JORGE MANOEL TEIXEIRA CARNEIRO (jorgemtc@iag.puc-rio.br)

Doutor em Administração de Empresas pela Coppead/UFRJ

Instituição de vinculação: Pontifícia Universidade Católica do Rio de Janeiro/IAG

Rio de Janeiro/RJ – Brasil

Áreas de interesse em pesquisa: Estratégia de Empresas e Negócios Internacionais.

LUÍS ANTÔNIO DA ROCHA DIB (dib@coppead.ufrj.br)

Doutor em Administração de Empresas pela Coppead/UFRJ

Instituição de vinculação: Coppead/Universidade Federal do Rio de Janeiro

Rio de Janeiro/RJ – Brasil

Áreas de interesse em pesquisa: Negócios Internacionais e Estratégia de Empresas.

Recebido em: 04/09/2010 • **Aprovado em:** 10/11/2011