

Paula Vicente e Elizabeth Reis¹

Sondagens na internet: como seleccionar os respondentes

Palavras-chave: Sondagens, Internet, Amostragem

Web surveys: how to select the respondents

Keywords: Surveys, Internet, Sampling

Resumo

As sondagens na Internet são cada vez mais populares na pesquisa de mercados pelas vantagens inquestionáveis de baixo custo, rapidez de recolha e processamento dos dados e potencialidades na construção dos questionários. A sua utilização fica no entanto restrita a contextos específicos de pesquisa onde a população-alvo coincide com a população da Internet. Ainda assim, nestes casos, o modo como é constituída a amostra é determinante da validade dos resultados. Neste artigo abordam-se os métodos de selecção das amostras em sondagens na Internet, discutem-se os pontos fracos de cada método e as estratégias que podem ser adoptadas para obter amostras representativas.

Abstract

Internet surveys are becoming increasingly popular in market research for the significant advantages of low cost, speed in data collecting and processing and technological enhancements in questionnaire design. Its application is however restricted to specific research problems, namely the ones when the target population of the survey matches the cybernauts population. But even in these cases, the method for sample selection is crucial for the validity of surveys' results. This paper reviews the sampling methods for internet surveys and discusses the weaknesses and strengths of each method. Strategies to assure representative samples are also discussed.

1. INTRODUÇÃO

As sondagens são tradicionalmente feitas através de entrevista telefónica, entrevista pessoal ou correio. O aparecimento da Internet veio alargar este leque de possibilidades prevendo-se que em 2007 as sondagens baseadas na tecnologia Internet representem 1/3 de todas as sondagens, a nível mundial; não é por isso improvável que no futuro a maioria das sondagens venha a ser feita on-line (Evans e Mathur, 2005), à semelhança do que aconteceu com o telefone que progressivamente conquistou terreno e destronou as sondagens por entrevista pessoal.

Os benefícios que a Internet oferece como meio de sondagem são claros: baixo custo, rapidez na recolha, processamento e análise dos dados, flexibilidade na concepção dos questionários com funcionalidades que os métodos tradicionais não permitem (Solomon, 2001). Além destes aspectos, a Internet torna a realização dos estudos de larga escala acessível a todos, pois qualquer pessoa/empresa com acesso à Internet e com capacidade

para trabalhar em ambiente Web pode criar um questionário que posteriormente envia por e-mail ou coloca num site, ficando imediatamente acessível a um número infundável de potenciais respondentes a custos muito inferiores aos conseguidos por outros métodos.

Embora evidenciando importantes pontos fortes, há um cepticismo persistente entre os utilizadores de sondagens quando confrontados com sondagens na WorldWideWeb; em muitos casos esse cepticismo até tem fundamento. As objecções mais frequentes são: (a) falta de representatividade da população em geral, sendo estas sondagens adequadas somente a populações muito específicas; (b) enviesamento de selecção nas amostras, pois normalmente é difícil, ou mesmo impossível, seleccionar aleatoriamente as amostras; (c) baixas taxa de resposta, por serem sondagens de auto-resposta, (d) problemas de hardware e software, que fazem com que nem todos os respondentes acedem à sondagem nas mesmas condições, (e) fraco controlo sobre o processo de

¹ UNIDE – Unidade de Investigação em Desenvolvimento Empresarial da IBS- ISCTE Business School

“entrevista”. As sondagens na Internet são também encaradas com desconfiança por razões de privacidade e segurança, bem como pelo receio generalizado de vírus e e-mails de spam (Umbach, 2004).

Como qualquer outro método, a Internet é adequada como meio de sondagem dependendo dos objectivos específicos e do propósito científico da investigação. As sondagens na Internet são mais eficientes (1) quando uma grande cobertura geográfica é procurada (e.g. sondagens multinacionais), (2) quando existem (e estão acessíveis) boas listagens da população que se pretende estudar, (3) quando se pretende fazer uma pesquisa exploratória ou experimental, ou (4) quando uma abordagem multimédia é desejável. Por exemplo, uma sondagem que tenha como objectivo determinar a reacção da população em geral a um novo canal televisivo, se realizada pela Internet deixará de fora todos os telespectadores que vêm televisão mas que não têm acesso à Internet; neste caso a sondagem teria um problema de cobertura pois através da Internet nem toda a população-alvo seria efectivamente estudada. Há casos em que a cobertura da população não é um problema. É o que acontece por exemplo com os quadros de empresas, os fornecedores ou clientes de um serviço, os utilizadores de um site ou de um serviço on-line, os estudantes de uma universidade, ou seja, domínios onde a utilização da Internet pode chegar aos 100%; embora estas sejam populações fechadas, a Internet é, do ponto de vista da cobertura, um bom meio para as sondar.

Para estudar populações alargadas a Internet tem algumas limitações. Em Portugal, a proporção de agregados familiares com ligação à Internet é de apenas 13% e a proporção de utilizadores da Internet em casa é somente de 18% (INE, 2001); a Internet deixa ainda de fora uma parte muito grande da população e está por isso ainda muito longe de ser uma boa solução para sondagens a agregados familiares ou à população geral. A Internet só é um bom ambiente para uma pesquisa se conseguir uma cobertura boa da população-alvo. Só depois desta condição assegurada vale a pena pensar no melhor desenho da sondagem tendo em conta os objectivos do estudo.

A Internet é um meio que não foi criado para fazer sondagens. A implementação de sondagens na Internet surgiu, aliás, pela mão de informáticos, que dominam muito bem a tecnologia informática da Web, mas que pouco (ou nada) sabem de sondagens. A facilidade e rapidez com que qualquer indivíduo, com razoáveis conhecimentos informáticos, monta uma sondagem na Internet justificam em grande parte a proliferação de sondagens na Internet. No entanto, cuidado, pois nem todos os procedimentos de recolha de informação que se encontram na Internet podem ser considerados sondagens. Há um aspecto metodológico que ajuda

a perceber se se está na presença de uma sondagem ou de uma outra qualquer actividade de recolha de informação com objectivos diversos de uma sondagem: trata-se do procedimento adoptado para a selecção dos respondentes. De facto, se a selecção dos respondentes não acontece segundo critérios de representatividade, a extrapolação dos resultados à população-alvo é questionável. A escolha do método de selecção dos respondentes é por isso crucial no desenho de uma boa sondagem na Internet (Dillman, 2000).

Neste artigo aborda-se o problema da selecção dos respondentes em sondagens na Internet, discutindo-se vários métodos de amostragem e apontando-se estratégias para uma boa prática na selecção das amostras. As sondagens na Internet podem assumir diversos formatos e os métodos de selecção dos respondentes são em grande parte determinados pelo formato da sondagem. Por este facto, o artigo começa com uma apresentação das várias formas de sondagem na Internet, abordando-se depois os principais métodos de selecção dos respondentes discutindo as limitações que cada um coloca à sondagem. Em Portugal, a utilização da Internet como meio de sondagem encontra-se ainda numa fase muito embrionária e, por esse facto, nem todos os tipos de sondagem que se descrevem estão igualmente generalizados, no entanto, merecem ser mencionados para melhor perceber as potencialidades que a Internet oferece em termos de desenho de sondagens.

2. FORMATOS DE SONDAEM NA INTERNET

As sondagens conduzidas na Internet podem assumir vários formatos que se distinguem entre si pela flexibilidade na concepção do questionário, pelo grau de controlo sobre a sondagem, pelo custo e pela duração da sondagem (Couper, 2000; Hewson et al, 2002; Manfreda et al, 2002). Com base nestes parâmetros, identificam-se essencialmente quatro formatos de sondagem na Internet: a) sondagens por e-mail, b) sondagens em HTML na Web, c) sondagens interactivas na Web; e d) sondagens por downloads. De forma sumária, caracteriza-se cada um dos formatos e referem-se as suas principais vantagens e desvantagens.

2.1. Sondagens por e-mail

Este formato é considerado a versão on-line de um questionário por correio. O respondente recebe o questionário na sua caixa de correio electrónico, consegue visualizar todo o questionário antes de responder, pode modificar as suas respostas em qualquer momento da sondagem e pode até alterar a formulação da questão para que ela reflecta melhor o que ele quer responder.

As formas mais elementares de sondagem por e-mail não incluem qualquer processamento automático das respostas, o que significa que a introdução das respostas numa base de dados informatizada é feita manualmente após a recepção dos e-mails com os questionários respondidos. Não há mecanismos interactivos de

controlo ou de teste da consistência das respostas, ao contrário do que acontece noutros formatos mais sofisticados.

As sondagens por e-mail são fáceis de conceber, são económicas e rápidas (entre o início e fim do estudo podem existir entre 3 horas a 3 dias). São idealmente aplicadas em ambientes organizativos quando se quer inquirir os trabalhadores, os clientes, os fornecedores ou qualquer outro tipo de “população interna” porque usualmente estes respondentes são capazes de respeitar as instruções de resposta ao questionário, acrescentar informação qualitativa relevante nas suas respostas, são capazes de cumprir as instruções de “salto” das questões, não alteram o texto das questões, ao contrário do que se verifica com populações “externas” e mais alargadas (Schaefer e Dillman, 1998).

2.2. Sondagens em HTML na Web

Estas sondagens assumem a forma de uma página HTML, mais ou menos longa, com questões para as quais a resposta é dada clicando em “botões” ou preenchendo caixas de texto. O respondente pode fazer o scrolling de toda a página (i.e. percorrê-la do início ao fim) e visualizar todo o questionário antes de responder.

Estas sondagens requerem capacidades acrescidas de programação, caso se pretenda transferir automaticamente para uma base de dados as respostas, à medida que os questionários são submetidos. Estas sondagens não têm verdadeiros mecanismos de controlo interactivos do questionário (i.e., não há “saltos” automáticos entre questões, o inquirido pode alterar a ordem por que responde às questões), mas são muito flexíveis em termos do design que pode ser conseguido no questionário. Podem utilizar-se efeitos gráficos, de som, de imagem, de animação ou outras formas multimédia para tornar o questionário mais atractivo.

Para estudos que se querem rápidos, o formato HTML pode ser uma boa solução pois a duração média destas sondagens varia entre 3 dias e 10 dias. O baixo custo é também uma vantagem, face a outras alternativas mais sofisticadas. São igualmente uma boa opção quando o encadeamento lógico das questões não é complexo. Há linguagens específicas para programação em HTML mas uma ferramenta de uso corrente acessível a (quase) todos é o Microsoft Front Page integrado no Microsoft Office (ambiente Windows).

2.3. Sondagens interactivas na Web

Esta forma de sondagem é como que uma versão em ambiente Web de um questionário em sistema CATI (Computer Assisted Telephone Interviews); as questões aparecem por ecrãs na Web, da mesma forma que apareceriam no posto de tele-sondagem do entrevistador. As ferramentas que permitem criar estas sondagens existem sob a forma de packages informáticos que o investigador

utiliza no seu computador pessoal. Embora alguns packages exijam qualificações especiais a nível de programação informática, há outros que são user friendly, o que coloca a concepção de uma sondagem deste tipo ao alcance de (quase) todos. Há naturalmente a necessidade de um investimento inicial na aquisição e aprendizagem de utilização do software.

O layout do questionário pode ser personalizado, segundo os requisitos do investigador, através da manipulação do formato das perguntas e respostas, do número de questões que surgem por ecrã, da escolha do padrão de fundo dos ecrãs ou da introdução de gráficos ou imagens (estáticos ou dinâmicos).

Os sistemas de controlo da administração do questionário e da consistência entre respostas desenvolvidos para as sondagens CATI (por exemplo, programar a ordem em que as questões surgem, automatizar o salto entre questões ou desenvolver rotinas de validação de consistência entre respostas) são também aplicáveis a estas sondagens.

As sondagens interactivas são caras e mais demoradas que as outras opções de sondagem já descritas, devido, sobretudo, ao custo e tempo envolvidos na programação da sondagem, que saem incrementados dependendo da complexidade da estrutura informática de dados que suporta a “personalização” criada. Estas sondagens podem durar de 7 dias a 14 dias.

2.4. Sondagens por downloads

Estas sondagens são descarregadas da Web e “corridas” no computador pessoal do respondente mediante a instalação prévia de software compatível, fornecido pelo investigador/empresa responsável pela sondagem; estas sondagens não são, por essa razão, verdadeiramente on-line. À semelhança das sondagens interactivas, as sondagens por downloads são construídas por ecrãs, que vão surgindo no computador pessoal do respondente à medida que este avança na resposta às questões. Todas as outras possibilidades em termos de desenho do questionário descritas para as sondagens interactivas são aqui igualmente possíveis. Quando o respondente “corre” a sondagem no seu computador é gerado, automaticamente, um ficheiro de dados que guarda as respostas dadas; o respondente deverá, posteriormente, fazer o upload deste ficheiro quando terminar o preenchimento do questionário.

As sondagens por downloads tendem a ser mais caras e demoradas do que as sondagens on-line. Em alguns casos requerem do respondente a capacidade para instalar o software que permite “correr” a sondagem e a capacidade para manusear correctamente o processo de upload do ficheiro de dados, o que, de alguma forma, pode limitar o perfil da população a estudar. Quando o software necessário à execução da sondagem tem de

ser descarregado para o computador do respondente pode haver desmotivação em participar quando o tempo envolvido nesta tarefa é demasiado (pode ser de apenas 20 minutos ou pode chegar a 2 horas!).

O facto do respondente trabalhar off-line origina com frequência o esquecimento no envio do ficheiro de dados o que na prática se traduz na perda de respondentes. Esta forma de sondagem é mais usada com painéis ou com grupos pré-recrutados de respondentes que têm o compromisso de regularmente comunicar com a empresa de sondagens. O tempo de duração de uma sondagem por downloads pode ir de 7 dias a 20 dias.

Existem vários programas informáticos para concepção de sondagens na Internet em formato interactivo; a referência de sites onde se pode encontrar informação sobre alguns desses softwares é apresentada após a bibliografia.

3. SELECÇÃO DOS RESPONDENTES

A sondagem é, por definição, um estudo de larga escala que se baseia numa amostra seleccionada da população que se pretende estudar. A selecção da amostra é, por isso, uma das etapas mais importantes na concepção da sondagem. A abordagem tradicional para descrever os métodos de amostragem nas sondagens em geral faz uma separação dos métodos em dois grupos: métodos aleatórios e métodos não aleatórios (Groves et al, 2004). No caso particular das sondagens na Internet, embora se possa também adoptar a dicotomia aleatório/não-aleatório para caracterizar os métodos de selecção dos respondentes, parece-nos mais adequado considerar a caracterização dos métodos segundo a fonte da amostra. Assim, nas sondagens pela Internet os respondentes podem ser seleccionados a partir de uma listagem de endereços de e-mail, podem ser escolhidos a partir de um painel pré-constituído de (potenciais) respondentes ou podem simplesmente ser “apanhados” para a sondagem ao serem interceptados quando navegam na Web. Sempre que se disponha de um registo que possa servir como base de sondagem, a selecção dos respondentes pode ser feita segundo um procedimento aleatório (simples, sistemático, estratificado ou de qualquer outro modo praticado em sondagens por outras vias) a partir dessa base. Quando não se dispõe de uma base de sondagem a selecção faz-se por conveniência. Se existir informação que o permita, podem introduzir-se quotas para controlar algumas características dos respondentes e dessa forma melhorar a representatividade da amostra.

Segundo a fonte da amostra, os métodos de selecção de respondentes em sondagens na Internet podem ser: a) auto-selecção, b) painéis voluntários de utilizadores da Internet, c) intercepção na Web, d) listagens de utilizadores da Internet

e, e) painéis pré-recrutados de utilizadores da Internet. A seguir, descrevem-se os métodos e discutem-se as vantagens e desvantagens de cada um. Como ficará patente, o método de selecção dos respondentes está muito associado ao formato da sondagem.

3.1. Auto-selecção

Uma amostra constituída por auto-selecção resulta exclusivamente da decisão do respondente em participar na sondagem. A sondagem é colocada em portais ou sites frequentemente visitados e está disponível para todos os que queriam responder; na prática, todos os que entram no site da sondagem podem integrar a amostra. Não há restrições de acesso e pouco ou nenhum controlo é feito sobre o número de vezes que um mesmo indivíduo responde. O convite para colaborar com a sondagem pode também ser divulgado por uma via que não a própria Internet, como sejam um jornal, a televisão ou o correio com a indicação URL da sondagem. No entanto, a decisão de responder é inteiramente do respondente.

Conceptualmente, a população-alvo é formada por todos os que visitam o site, mas não há uma base de sondagem que permita referenciar de modo mais preciso a população que está a ser estudada. Não há, no entanto, como garantir que os que se auto-seleccionam para a sondagem representam aquela população. As amostras constituídas por auto-selecção são normalmente enviesadas nas variáveis substantivas pois os que são mais sensíveis ao tema em estudo são os que têm mais probabilidade de responder e não há garantia que estes indivíduos sejam representativos de um conjunto mais alargado de indivíduos. Em rigor, os resultados de uma sondagem baseada numa amostra constituída por auto-selecção são apenas representativos do grupo que respondeu e não devem ser extrapolados para outras realidades.

3.2. Painéis voluntários de utilizadores da Internet

Os painéis de utilizadores da Internet formam-se antes da realização de qualquer sondagem e são constituídos por cibernautas que voluntariamente aderem ao apelo para colaborar em sondagens. Estes apelos podem ser colocados em sites ou portais muito visitados, ou mesmo no site da empresa de sondagens. No momento em que alguém se regista é recolhida informação demográfica a seu respeito, o que permite criar uma base de dados de potenciais respondentes que funcionará como base de sondagem em futuras sondagens.

A selecção dos respondentes para uma sondagem específica a partir do painel constituído pode ser aleatória ou basear-se em quotas. O acesso à sondagem é feito apenas por convite directo ao respondente e controlado por identificadores de e-mail e passwords.

Embora a selecção dos respondentes possa ser aleatória e a amostra possa ser desenhada de forma a controlar vários atributos demográficos dos respondentes, o facto de o painel inicialmente criado ser constituído por voluntários limita o âmbito das conclusões destas sondagens. Um painel de voluntários não é representativo da população de utilizadores da Internet e menos ainda da população em geral, por isso, uma sondagem baseada numa amostra constituída a partir de um painel de voluntários é, quando muito, representativa apenas da população do painel.

3.3. Intercepção na Web

A selecção por intercepção na Web acontece quando um cibernauta, ao entrar num determinado site, é interceptado e convidado a responder à sondagem; mensagens publicitárias mostradas em banners conduzem o cibernauta ao site da sondagem. A cadência com que estas mensagens (ou seja, os convites) surgem no site pode ser previamente definida para reproduzir um esquema de selecção semelhante ao que se pratica nas sondagens à boca das urnas, ou seja, implementa-se um procedimento de selecção sistemática para “convidar” cada n-ésimo visitante do site a participar na sondagem.

A selecção de respondentes por intercepção na Web é frequentemente adoptada em sondagens sobre a satisfação dos consumidores on-line de um serviço, ou em sondagens de avaliação da qualidade de um site. A população-alvo é restrita aos visitantes ou utilizadores do site, mas ainda assim, a selecção por intercepção não garante que todos os elementos desta população tenham alguma probabilidade de ser seleccionados, uma vez que não existe uma base de sondagem fixa de onde previamente se possam seleccionar os respondentes.

Para evitar que a amostra fique composta pelos visitantes (mais) frequentes do site pode fazer-se um screening inicial de algumas características dos cibernautas nomeadamente a sua frequência de visita ao site, se já participou numa sondagem através daquele mesmo site e em que data. Para evitar que um mesmo indivíduo responda à sondagem mais do que uma vez podem introduzir-se cookies, ainda que este procedimento seja eticamente questionável.

A selecção por intercepção permite aceder a milhões de participantes, pois todos os que navegam e visitam o site são potenciais respondentes. No entanto não é possível quantificar ou referenciar de forma mais exacta a população estudada, pelo que os resultados de uma sondagem feita nestes moldes traduzem apenas a opinião dos que responderam e não devem ser extrapolados para além disso.

3.4. Listagens de utilizadores da Internet

A selecção a partir de listagens de utilizadores da Internet é, na prática, uma selecção a partir de listagens de endereços electrónicos. Estas listas funcionam como bases de sondagem e a selecção dos respondentes pode inclusivamente ser aleatória. O convite para a sondagem é enviado por e-mail para os indivíduos seleccionados.

As sondagens em contexto organizacional ou as sondagens dirigidas a utilizadores da Internet adoptam este procedimento de amostragem sempre que existam listas de endereços de e-mail. A população-alvo está referenciada pela listagem de endereços e, se a amostra for seleccionada aleatoriamente, é possível extrapolar os resultados da sondagem a esta população.

3.5. Painéis pré-recrutados de utilizadores da Internet

Os painéis pré-recrutados de utilizadores da Internet são formados por indivíduos que se disponibilizam para participar em sondagens, mas, ao contrário do que acontece nos painéis de voluntários, nestes painéis os potenciais painelistas são contactados telefonicamente. A selecção do número de telefone para este contacto é feita aleatoriamente a partir de uma lista telefónica, ou através da marcação aleatória de números de telefone. Este primeiro contacto serve para recolher alguma informação sócio-demográfica sobre o indivíduo, nomeadamente se tem acesso/é utilizador da Internet e se quer integrar o painel para colaboração em futuras sondagens.

Os painéis pré-recrutados de utilizadores da Internet funcionam como base de sondagem e permitem a selecção aleatória de respondentes em cada sondagem concreta. A informação sócio-demográfica obtida de cada respondente permite introduzir variáveis de controlo no desenho da amostra.

Cada painalista tem um número de identificação pessoal ou uma password que o identifica perante a empresa (ou a entidade que faz as sondagens) em cada sondagem. A solicitação para participar em cada sondagem é enviada para o e-mail do respondente e o acesso à sondagem é controlado por ID's integrados na URL da sondagem, para assegurar que só os indivíduos seleccionados respondem e que o fazem apenas uma vez.

A qualidade das amostras constituídas a partir de um painel pré-recrutado depende em larga medida de como foi constituído o painel. O painel é constituído com referência a uma população-alvo. Se a escolha dos painelistas tiver em conta critérios que asseguram a representatividade dessa população, se os procedimentos de manutenção do painel ao longo do tempo assegurarem que em cada momento o painel é uma boa representação da população-alvo e se, em paralelo, a selecção dos

respondentes para cada sondagem acontece segundo critérios de representatividade, então os resultados das sondagens baseadas em painéis pré-recrutados são extrapoláveis à população-alvo.

4. GARANTIR A QUALIDADE DAS AMOSTRAS – O QUE PODE SER FEITO

A qualidade de uma sondagem na Internet é, em grande medida, determinada pela qualidade da amostra que lhe está subjacente, mas é de notar que o juízo sobre a qualidade de uma amostra deve ser enquadrado pelo objectivo da sondagem e pelo âmbito de generalização que se pretende dar aos resultados. Uma amostra constituída por intercepção na Web pode ser adequada numa sondagem destinada a conhecer a opinião dos cibernautas sobre um determinado site, mas é desajustada numa sondagem em que se pretende conhecer a preferência dos consumidores em geral por uma nova bebida light.

Independentemente da especificidade de cada sondagem, é consensual que a amostra deve ser representativa da realidade que se propõe estudar e que as conclusões da sondagem não podem ser extrapoladas para além da realidade que a amostra espelha. Estes dois aspectos são naturalmente condicionados pela fonte da amostra e pelo procedimento propriamente dito de selecção dos respondentes. As seguintes técnicas podem ser usadas para assegurar a representatividade da amostra

4.1. Introduzir variáveis de controlo na selecção dos respondentes

Os respondentes de uma sondagem na Internet podem ser conduzidos até à sondagem através de um e-mail que é recebido e no qual está o link para o site da sondagem. Em alternativa, o respondente pode ser interceptado enquanto navega na Web, depois de um screening que o identifica como potencial respondente. Num caso ou noutro, se os convites para a sondagem forem enviados a uma amostra equilibrada a probabilidade de obter uma amostra equilibrada é maior. Uma amostra equilibrada é uma amostra que reproduz a estrutura da população em características que são importantes para o estudo. Na prática, desenhar uma amostra equilibrada passa por introduzir variáveis de controlo que funcionam como estratos ou quotas, na selecção dos respondentes.

Para garantir que as quotas são cumpridas, é necessário um acompanhamento continuado da sondagem para fazer a verificação do preenchimento das quotas à medida que os questionários respondidos são recebidos. O envio de lembretes insistindo na resposta é importante para assegurar o cumprimento das quotas estabelecidas.

4.2. Aleatoriedade na selecção dos respondentes

A selecção aleatória dos respondentes deve ser sempre preferida

à selecção não-aleatória, pois esta corre o risco de originar uma amostra enviesada. Há sondagens na Internet em que não é possível escapar à escolha não-aleatória, como sejam o caso da auto-selecção ou da selecção por intercepção. Nos casos em que exista um registo da população que possa funcionar como base de sondagem, a selecção dos respondentes pode e deve ser aleatória, mas é de notar que a possibilidade de extrapolar os resultados da sondagem além da amostra depende da qualidade da base de sondagem.

4.3. Trabalhar com boas bases de sondagem

A qualidade de uma base de sondagem é determinada pela forma como é constituída. No caso de listagens de endereços de e-mail é importante a exaustividade, i.e., que a listagem represente todos os indivíduos da população-alvo. No caso dos painéis é importante que o painel seja constituído de forma a ser representativo da população-alvo. Os painéis de voluntários têm uma qualidade questionável, uma vez que apenas a vontade do indivíduo em colaborar determina a sua pertença ao painel, i.e., não existe uma selecção independente da vontade do indivíduo. Nos painéis pré-recrutados, há um procedimento aleatório que determina o contacto com o indivíduo; se na formação do painel for ainda considerada a distribuição da população em características relevantes para a sondagem, então pode falar-se em painel representativo da população.

Há painéis onde um mesmo conjunto de indivíduos participa regularmente em sondagens; são painéis de difícil manutenção pois são exigentes para o respondente que periodicamente se vê “obrigado” a colaborar, mas que permitem taxas de resposta que podem ir de 35% a 75%. Outros painéis são compostos por indivíduos que, tendo participado numa sondagem em determinado momento, não rejeitaram a possibilidade de serem novamente contactados para futuras sondagens, mas nestes painéis as taxas de resposta são apenas da ordem dos 5 a 10%.

4.4. Ponderação da amostra

Ponderar a amostra consiste em atribuir um factor multiplicativo a cada elemento da amostra obtida, alterando o seu peso relativo na análise dos dados. Alguns tipos de ponderação visam equilibrar as probabilidades de selecção. Por exemplo, numa sondagem por auto-selecção, os cibernautas que estão mais tempo on-line têm maior probabilidade de aparecer na amostra do que os que estão menos tempo, pelo que cada respondente deveria ser ponderado pelo inverso do tempo dispendido on-line. Noutros casos, a ponderação pode visar equilibrar a distribuição da amostra em uma ou mais variáveis. Em qualquer dos casos, o cálculo de ponderadores requer informação relativa à(s) variável(is) a considerar na definição do esquema de ponderação e nem sempre essa informação está disponível nas populações da Internet.

Se a amostra é aleatória mas apresenta enviesamento em algumas variáveis (o que pode ser consequência das não-respostas), a ponderação pode ser uma opção para reequilibrar a amostra obtida. Quando o enviesamento é grande, a ponderação é questionável, pois a aplicação de ponderadores não substitui um bom procedimento de selecção, da mesma forma que não consegue corrigir a amostra se o procedimento de selecção foi inapropriado (Manfreda et al, 1999). A ponderação da amostra é, pois, uma solução de recurso capaz de atenuar pequenos enviesamentos, mas incapaz de corrigir uma amostra mal desenhada, além, de que é exequível somente quando existem dados populacionais sobre as variáveis de ponderação.

O que acontece à qualidade da amostra se as taxas de resposta forem baixas? Neste caso, a validade dos resultados pode ficar comprometida. Qualquer que tenha sido o procedimento de amostragem adoptado, uma baixa taxa de respostas compromete sempre a representatividade da amostra, por isso, conseguir manter a taxa de não-respostas em valores baixos é outro desafio que se coloca no planeamento e na execução de uma sondagem na Internet; mas o problema das não-respostas é igualmente vasto e merece ser tratado autonomamente do problema da selecção dos respondentes.

5. COMENTÁRIOS FINAIS

O valor da Internet como ferramenta de pesquisa é claro, com vantagens inquestionáveis de rapidez, baixo custo e potencialidades de construção dos questionários sobre os métodos tradicionais de sondagem. Mas as sondagens na Internet merecem ser avaliadas e debatidas também e, sobretudo, em termos dos seus aspectos metodológicos, nomeadamente o tipo de populações que conseguem cobrir, a forma como se constituem as amostras e a taxa de respostas assegurada.

As sondagens na Internet podem assumir vários formatos e os últimos desenvolvimentos são mais tecnológicos do que metodológicos. No entanto, dispor das ferramentas logísticas para lidar com a Internet não é, por si só, suficiente para montar uma sondagem de qualidade - os aspectos metodológicos devem igualmente ser acautelados na medida em que é a componente metodológica da sondagem que determina a validade dos resultados. A escolha de um "fornecedor de sondagens" não deve, pois, ser baseada somente nos serviços técnicos que oferece, como o desenho do questionário ou o processamento e análise dos dados, mas sobretudo na competência em tratar com rigor a metodologia da sondagem e na capacidade em expor os detalhes necessários à compreensão clara sobre a validade dos dados recolhidos via Internet.

REFERÊNCIAS

- COUPER, Mick, (2000), "Web surveys: a review of issues and approaches", *Public Opinion Quarterly*, 64, 464-494.
- DILLMAN, Don, (2000), *Mail and internet surveys: the tailored design method* (2nd Edition), John Wiley & Sons, New York.
- EVANS, Joel; MATHUR, Anil, (2005), "The value of on-line surveys", *Internet Research*, 15, 195-219.
- GROVES, Robert; FOWLER, Floyd; COUPER, Mick; LEPKOWSKI, James; SINGER, Eleanor; TOURANGEAU, Roger, (2004), *Survey methodology*, John Wiley & Sons, New York.
- HEWSON, Claire; YULE, Peter; LAURENT, Dianna; VOGEL, Carl, (2002), *Internet research methods: a practical guide for the social and behavioural sciences*, Sage Publications, London.
- INE, (2001), *Inquérito à utilização das tecnologias de informação*, Instituto Nacional de Estatística, Portugal.
- MANFREDA, Lozar; BATAGELJ, Zenel; VEHOVAR, Vasja, (2002), "Design of Web survey questionnaires: three basic experiments", *Journal of Computed-Mediated Communications*, 7, 3.
- SCHAEFER, David; DILLMAN, Don, (1998), "Development of a standard Email methodology", *Public Opinion Quarterly*, 62, 378-397.
- SOLOMON, David, (2001), "Conducting Web-based surveys", *Practical Assessment, Research & Evaluation*, 7, 1-6.
- UMBACH, Paul, (2004), "Web surveys: best practices", *New Directions for Institutional Research*, 121, 23-38.

ENDEREÇOS URL DE SOFTWARE PARA INQUÉRITOS ON-LINE

- <http://www.hostedsurvey.com>
- <http://www.pse.pt>
- <http://www.quask.com>
- <http://www.vanguardsw.com>