14

Sampling Methods for Online Surveys

Ronald D. Fricker, Jr

INTRODUCTION

In the context of conducting surveys or collecting data, sampling is the selection of a subset of a larger population to survey. This chapter focuses on sampling methods for web and e-mail surveys, which taken together we call 'online' surveys. In our discussion we will frequently compare sampling methods for online surveys to various types of non-online surveys, such as those conducted by postal mail and telephone, which in the aggregate we refer to as 'traditional' surveys.

The chapter begins with a general overview of sampling. Since there are many fine textbooks on the mechanics and mathematics of sampling, we restrict our discussion to the main ideas that are necessary to ground our discussion on sampling for online surveys. Readers already well versed in the fundamentals of survey sampling may wish to proceed directly to the section on Sampling Methods for online Surveys.

WHY SAMPLE?

Surveys are conducted to gather information about a population. Sometimes the survey is conducted as a *census*, where the goal is to survey every unit in the population. However, it is frequently impractical or impossible to survey an entire population, perhaps owing to either cost constraints or some other practical constraint, such as that it may not be possible to identify all the members of the population.

An alternative to conducting a census is to select a sample from the population and survey only those sampled units. As shown in Figure 14.1, the idea is to draw a sample from the population and use data collected from the sample to infer information about the entire population. To conduct *statistical inference* (i.e., to be able to make quantitative statements about the unobserved population statistic), the sample must

be drawn in such a fashion that one can be confident that the sample is representative of the population and that one can both calculate appropriate sample statistics and estimate their standard errors. To achieve these goals, as will be discussed in this chapter, one must use a probability-based sampling methodology.
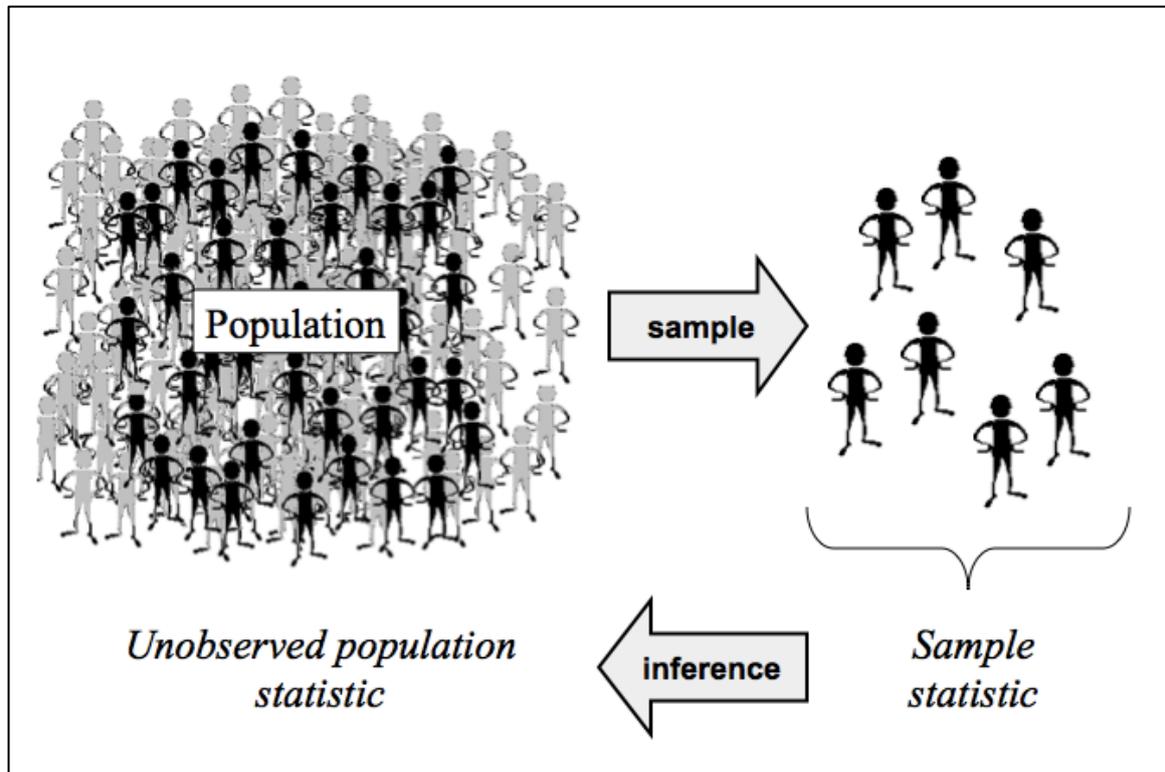

Figure 14.1 An illustration of sampling. When it is impossible or infeasible to observe a population statistic directly, data from a sample appropriately drawn from the population can be used to infer information about the population. (Source: author)

A survey administered to a sample can have a number of advantages over a census, including:

- lower cost

- less effort to administer

- better response rates

- greater accuracy.

The advantages of lower cost and less effort are obvious: keeping all else constant, reducing the number of surveys should cost less and take less effort to field and analyze. However, that a survey based on a sample rather than a census can give better response rates and greater accuracy is less obvious. Yet, greater survey accuracy *can* result when the sampling error is more than offset by a decrease in nonresponse and other

biases, perhaps due to increased response rates. That is, for a fixed level of effort (or funding), a sample allows the surveying organization to put more effort into maximizing responses from those surveyed, perhaps via more effort invested in survey design and pre-testing, or perhaps via more detailed non-response follow-up.

What does all of this have to do with online surveys? Before the Internet, large surveys were generally expensive to administer and hence survey professionals gave careful thought to how to best conduct a survey in order to maximize information accuracy while minimizing costs. However, the Internet now provides easy access to a plethora of inexpensive survey software, as well as to millions of potential survey respondents, and it has lowered other costs and barriers to surveying. While this is good news for survey researchers, these same factors have also facilitated a proliferation of bad survey research practice.

For example, in an online survey the marginal cost of collecting additional data can be virtually zero. At first blush, this seems to be an attractive argument in favor of attempting to conduct censuses, or for simply surveying large numbers of individuals without regard to how the individuals are recruited into the sample. And, in fact, these approaches are being used more frequently with online surveys, without much thought being given to alternative sampling strategies or to the potential impact such choices have on the accuracy of the survey results. The result is a proliferation of poorly conducted 'censuses' and surveys based on large convenience samples that are likely to yield less accurate information than a well-conducted survey of a smaller sample.

Conducting surveys, as in all forms of data collection, requires making compromises. Specifically, there are almost always trade-offs to be made between the amount of data that can be collected and the accuracy of the data collected. Hence, it is critical for researchers to have a firm grasp of the trade-offs they implicitly or explicitly make when choosing a sampling method for collecting their data.

AN OVERVIEW OF SAMPLING

There are many ways to draw samples from a population – and there are also many ways that sampling can go awry. We intuitively think of a good sample as one that is representative of the population from which the sample has been drawn. By 'representative' we do not necessarily mean the sample matches the

population in terms of observable characteristics, but rather that the results from the data we collect from the sample are consistent with the results we would have obtained if we had collected data on the entire population.

Of course, the phrase 'consistent with' is vague and, if this was an exposition of the mathematics of sampling, would require a precise definition. However, we will not cover the details of survey sampling here.[1] Rather, in this section we will describe the various sampling methods and discuss the main issues in characterizing the accuracy of a survey, with a particular focus on terminology and definitions, in order that we can put the subsequent discussion about online surveys in an appropriate context.

Sources of error in surveys

The primary purpose of a survey is to gather information about a population. However, even when a survey is conducted as a census, the results can be affected by several sources of error. A good survey design seeks to reduce all types of error – not only the sampling error arising from surveying a sample of the population. Table 14.1 below lists the four general categories of survey error as presented and defined in Groves (1989) as part of his 'Total Survey Error' approach.

Errors of coverage occur when some part of the population cannot be included in the sample. To be precise, Groves specifies three different populations:

1. The population of inference is the population that the researcher ultimately intends to draw conclusions about.

2. The target population is the population of inference less various groups that the researcher has chosen to disregard.

3. The frame population is that portion of the target population which the survey materials or devices delimit, identify, and subsequently allow access to (Wright and Tsao, 1983).

The *survey sample* then consists of those members of the sampling frame who are chosen to be surveyed, and *coverage error* is the difference between the frame population and the population of inference.

The two most common approaches to reducing coverage error are:

- obtaining as complete a sampling frame as possible (or employing a *frameless sampling strategy* in which most or all of the target population has a positive chance of being sampled);

- *post-stratifying* to weight the survey sample to match the population of inference on some observed key characteristics.

Sampling error arises when a sample of the target population is surveyed. It results from the fact that different samples will generate different survey data. Roughly speaking, assuming a random sample, sampling error is reduced by increasing the sample size.

Nonresponse errors occur when data is not collected on either entire respondents (*unit nonresponse*) or individual survey questions (*item nonresponse*). Groves (1989) calls nonresponse 'an error of nonobservation'. The response rate, which is the ratio of the number of survey respondents to the number sampled, is often taken as a measure of how well the survey results can be generalized. Higher response rates are taken to imply a lower likelihood of *nonresponse bias*.

Measurement error arises when the survey response differs from the 'true' response. For example, respondents may not answer sensitive questions honestly for a variety of reasons, or respondents may misinterpret or make errors in answering questions. Measurement error is reduced in a variety of ways, including careful testing and revision of the survey instrument and questions, choice of survey mode or modes, etc.

Table 14.1 Sources of survey error according to Groves (1989)

Type of error	Definition
Coverage	'…the failure to give any chance of sample selection to some persons in the population'.
Sampling	'…heterogeneity on the survey measure among persons in the population'.
Nonresponse	'…the failure to collect data on all persons in the sample'.
Measurement	'…inaccuracies in responses recorded on the survey instruments'.

Sampling methods

Survey sampling can be grouped into two broad categories: *probability-based sampling* (also loosely called 'random sampling') and *non-probability sampling*. A probability-based sample is one in which the respondents are selected using some sort of probabilistic mechanism, and where the probability with which

every member of the frame population could have been selected into the sample is known. The sampling probabilities do not necessarily have to be equal for each member of the sampling frame.

Types of probability sample include:

- *Simple random sampling* (SRS) is a method in which any two groups of equal size in the population are equally likely to be selected. Mathematically, simple random sampling selects n units out of a population of size N such that every sample of size n has an equal chance of being drawn.

- *Stratified random sampling* involves splitting the population up into non-overlapping strata which are then separately sampled. It is useful when the population is comprised of a number of homogeneous groups. In these cases, it can be either practically or statistically advantageous (or both) to first stratify the population into the homogeneous groups and then use SRS to draw samples from each group.

- *Cluster sampling* occurs when the natural sampling unit is a group or cluster of individual units. For example, in surveys of Internet users it is sometimes useful or convenient to first sample by discussion groups or Internet domains, and then to sample individual users within the groups or domains. In most (offline) face-to-face surveys for which no sampling frame exists, areal cluster sampling is used in which interviewers are sent to a location and then they sample some number of units that are in close geographic proximity.

- *Systematic sampling* is the selection of every k^{th} element from a sampling frame or from a sequential stream of potential respondents. Systematic sampling has the advantage that a sampling frame does not need to be assembled beforehand. In terms of Internet surveying, for example, systematic sampling can be used to sample sequential visitors to a website. The resulting sample is considered to be a probability sample as long as the sampling interval does not coincide with a pattern in the sequence being sampled and a random starting point is chosen.

There are important analytical and practical considerations associated with how one draws and subsequently analyzes the results from each of these types of probability-based sampling schemes, but space limitations preclude covering them here. Readers interested in such details should consult texts such as Kish (1965), Cochran (1977), Fink (2003), or Fowler and Floyd (2002).

Non-probability samples, sometimes called *convenience samples*, occur when either the probability that every unit or respondent included in the sample cannot be determined or it is left up to each individual to

choose to participate in the survey. For probability samples, the surveyor selects the sample using some probabilistic mechanism and the individuals in the population have no control over this process. In contrast, for example, a web survey may simply be posted on a website where it is left up to those browsing through the site to decide to participate in the survey ('opt in') or not. As the name implies, such non-probability samples are often used because it is somehow convenient to do so.

While in a probability-based survey participants can choose not to participate in the survey ('opt out'), rigorous surveys seek to minimize the number who decide not to participate (i.e., nonresponse). In both cases it is possible to have bias, but in non-probability surveys the bias has the potential to be much greater, since it is likely that those who opt in are not representative of the general population. Furthermore, in non-probability surveys there is often no way to assess the potential magnitude of the bias, since there is generally no information on those who chose *not* to opt in.

Non-probability-based samples often require much less time and effort, and thus usually are less costly to generate, but generally they do not support formal statistical inference. However, non-probability-based samples can be useful for research in other ways. For example, early in the course of research, responses from a convenience sample might be useful in developing hypotheses. Responses from convenience samples might also be useful for identifying issues, defining ranges of alternatives, or collecting other sorts of non-inferential data. For a detailed discussion on the application of various types of non-probability-based sampling method to qualitative research, see Patton (2002).

Specific types of non-probability samples include the following.

- *Quota sampling* requires the survey researcher only to specify quotas for the desired number of respondents with certain characteristics. The actual selection of respondents is then left up to the survey interviewers who must match the quotas. Because the choice of respondents is left up to the survey interviewers, subtle biases may creep into the selection of the sample (see, for example, the Historical Survey Gaffes section).

- *Snowball sampling* (also known as *respondent driven sampling*) is often used when the desired sample characteristic is so rare that it is extremely difficult or prohibitively expensive to locate a sufficiently large number of respondents by other means (such as simple random sampling). Snowball sampling relies on referrals from initial respondents to generate additional respondents.

While this technique can dramatically lower search costs, it comes at the expense of introducing bias because the technique itself substantially increases the likelihood that the sample will not be representative of the population.

- *Judgement sampling* is a type of convenience sampling in which the researcher selects the sample based on his or her judgement. For example, a researcher may decide to draw the entire random sample from one 'representative' Internet-user community, even though the population of inference includes all Internet users. Judgment sampling can also be applied in even less structured ways without the application of any random sampling.

Bias versus variance

If a sample is systematically not representative of the population of inference in some way, then the resulting analysis is likely to be biased. For example, results from a survey of Internet users about personal computer usage are unlikely to accurately quantify computer usage in the general population simply because the sample is comprised only of those who use computers. Furthermore, it is important to recognize that taking larger samples will not correct for bias, nor is a large sample evidence of a lack of bias. For example, an estimate of average computer usage based on a sample of Internet users will likely overestimate the average usage in the general population *regardless of how many Internet users are surveyed*.

Randomization, meaning randomly selecting respondents from the population of interest, is used to minimize the chance of bias. The idea is that by randomly choosing potential survey respondents *from the entire population* the sample is likely to 'look like' the population, even in terms of those characteristics that cannot be observed or known. This latter point is worth emphasizing. Probability samples mitigate the chance of sampling bias in both observable *and unobservable* characteristics.

Variance, on the other hand, is simply a measure of variation in the observed data. It is used to calculate the *standard error* of a statistic, which is a measure of the variability of the statistic. The precision of statistical estimates drawn via probabilistic sampling mechanisms is improved by larger sample sizes because (all else held constant) larger samples sizes result in smaller standard errors.

Some important sources of bias

Bias can creep into survey results in many different ways. In the absence of significant nonresponse, probability-based sampling is the best way to minimize the possibility of bias. Convenience sampling, on

the other hand, is generally assumed to have a higher likelihood of generating a biased sample. However, even with randomization, surveys of and about people may be subject to other kinds of bias. For example, respondents may be inclined to overstate or understate certain things ('sensitivity bias'), particularly with socially delicate questions (such as questions about income or sexual orientation, for example). Here we just focus on some of the more common sources of bias related to sampling.

- *Frame coverage* bias occurs when the sampling frame misses some important part of the population. For example, an e-mail survey using a list of e-mail addresses will miss those without an e-mail address.

- *Selection bias* is an error in how the individual or units are chosen to participate in the survey. It can occur, for example, if survey participation depends on the respondents having access to particular equipment, such as online surveys that miss those without Internet access.

- *Size bias* occurs when some units have a greater chance of being selected than others. For example, in a systematic sample of website visitors, frequent site visitors are more likely to get selected into the sample than those that do not. In a similar vein, when selecting from a frame consisting of e-mail addresses, individuals with multiple e-mail addresses would have a higher chance of being selected into a sample.

- *Nonresponse bias* occurs if those who refuse to answer the survey are somehow systematically different from those who do answer it.

Historical survey gaffes

A famous example of a survey that reached exactly the wrong inferential conclusion as a result of bias, in this case frame coverage and nonresponse bias, is the 'Literary Digest' poll in the 1936 United States presidential election. As described in Squires (1988), for their survey 'Literary Digest' assembled a sampling frame from telephone numbers and automobile registration lists. While using telephone numbers today might result in a fairly representative sample of the population, in 1936 only one in four households had a telephone and those were the more well-to-do. Compounding this, automobile registration lists only further skewed the frame towards individuals with higher incomes.

'Literary Digest' mailed 10 million straw-vote ballots, of which 2.3 million were returned, an impressively large number, but it represented less than a 25 per cent response rate. Based on the poll data,

'Literary Digest' predicted that Alfred Landon would beat Franklin Roosevelt 55 per cent to 41 per cent. In fact, Roosevelt beat Landon by 61 per cent to 37 per cent. This was the largest error ever made by a major poll and may have contributed to 'Literary Digest's' demise in 1938.

Gallup, however, called the 1936 presidential election correctly, even though he used significantly less data. But even Gallup, a pioneer in modern survey methods, didn't always get it right. In the 1948 United States presidential election between Harry S. Truman and Thomas E. Dewey, Gallup used a quota sampling method in which each pollster was given a set of quotas of types of people to interview, based on demographics. While that seemed reasonable at the time, the survey interviewers, for whatever conscious or subconscious reason, were biased towards interviewing Republicans more often than Democrats. As a result, Gallup predicted a Dewey win of 49.5 per cent to 44.5 per cent but almost the opposite occurred, with Truman beating Dewey with 49.5 per cent of the popular vote to Dewey's 45.1 per cent (a difference of almost 2.2 million votes).[2] Poor electoral forecasts for this election, not just by Gallup (and Gallup's mistaken forecast was due to more than quota sampling), are aptly summed up in the famous "Dewey Defeats Truman" photo in Figure 14.2.

Figure 14.2 Famous "Dewey Defeats Truman" photo. (Source: Library of Congress, 2012)

SAMPLING METHODS FOR INTERNET-BASED SURVEYS

This section describes specific online survey approaches and the sampling methods that are applicable to each. We concentrate on differentiating whether particular sampling methods and their associated surveys allow for generalization of survey results to populations of inference or not, providing examples of some surveys that were done appropriately and well, and others that were less so. Examples that fall into the latter category should not be taken as a condemnation of a particular survey or sampling method, but rather as illustrations of inappropriate application, execution, analysis, etc. Couper (2000: 465–466) perhaps said it best,

> Any critique of a particular Web survey approach must be done in the context of its intended purpose and the claims it makes. Glorifying or condemning an entire approach to survey data collection should not be done on the basis of a single implementation, nor should all Web surveys be treated as equal.

Furthermore, as we previously discussed, simply because a particular method does not allow for generalizing beyond the sample does not imply that the methods and resulting data are not useful in other research contexts.

Similar to Couper (2000), Table 14.2 lists the most common probability and non-probability sampling methods, and indicates which online survey mode or modes may be used with each method. For example, it is possible to conduct both web and e-mail surveys using a list-based sampling frame methodology. Conversely, while it is feasible to conduct an entertainment poll by e-mail, virtually all such polls are conducted via web surveys.

Table 14.2 Types of online survey and associated sampling methods

Sampling method	Web	E-mail
Probability-based		
Surveys using a list-based sampling frame	✓	✓
Surveys using non-list-based random sampling	✓	✓
Intercept (pop-up) surveys	✓	
Mixed-mode surveys with online option	✓	✓
Pre-recruited panel surveys	✓	✓
Non-probability		
Entertainment polls	✓	
Unrestricted self-selected surveys	✓	
Surveys using 'harvested' e-mail lists (and data)	✓	✓
Opt-in panels (volunteer or paid)	✓	

Surveys using a list-based sampling frame

Sampling for online surveys using a list-based sampling frame can be conducted just as one would for a traditional survey using a sampling frame. Simple random sampling in this situation is straightforward to implement and requires nothing more than contact information (generally an e-mail address for an online survey) on each unit in the sampling frame. Of course, though only contact information is required to field the survey, having additional information about each unit in the sampling frame is desirable to assess (and perhaps adjust for) nonresponse effects.

While online surveys using list-based sampling frames can be conducted either via the web or by e-mail, if an all-electronic approach is preferred the invitation to take the survey will almost always be made via e-mail. And, because e-mail lists of general populations are generally not available, this survey approach is most applicable to large homogeneous groups for which a sampling frame with e-mail addresses can be assembled (for example, universities, government organizations, large corporations, etc). Couper (2000) calls these 'list-based samples of high-coverage populations'.

In more complicated sampling schemes, such as a stratified sampling, auxiliary information about each

unit, such as membership in the relevant strata, must be available and linked to the unit's contact information. And more complicated multi-stage and cluster sampling schemes can be difficult or even impossible to implement for online surveys. First, to implement without having to directly contact respondents will likely require significant auxiliary data, which is unlikely to be available except in the case of specialized populations. Second, if offline contact is required, then the researchers are likely to have to resort to the telephone or mail in order to ensure that sufficient coverage and response rates are achieved.

An example of a multi-stage sampling procedure, used for an online survey of real estate journalists for which no sampling frame existed, is reported by Jackob et al. (2005). For this study, the researchers first assembled a list of publications that would have journalists relevant to the study. From this list a stratified random sample of publications was drawn, separately for each of five European countries. They then contacted the managing editor at each sampled publication and obtained the necessary contact information on all of the journalists that were 'occupied with real-estate issues'. All of the journalists identified by the managing editors were then solicited to participate in a web survey. Jackob et al. (2005) concluded that it 'takes a lot of effort especially during the phase of preparation and planning' to assemble the necessary data and then to conduct an online survey using a multi-stage sampling methodology.

Surveys using non-list-based random sampling

Non-list-based random sampling methods allow for the selection of a probability-based sample without the need to actually enumerate a sampling frame. With traditional surveys, random digit dialing (RDD) is a non-list-based random sampling method that is used mainly for telephone surveys.

There is no equivalent of RDD for online surveys. For example, it is not possible (practically speaking) to generate random e-mail addresses (see the Issues and Challenges in Online Survey Sampling section). Hence, with the exception of intercept surveys, online surveys requiring non-list-based random sampling depend on contacting potential respondents via some traditional means such as RDD, which introduces other complications and costs. For example, surveyors must either screen potential respondents to ensure they have Internet access or field a survey with multiple response modes. Surveys with multiple response modes introduce further complications, both in terms of fielding complexity and possible mode effects (again, see the Issues and Challenges in Online Survey Sampling section).

Intercept surveys

Intercept surveys on the web are pop-up surveys that frequently use systematic sampling for every k^{th} visitor to a website or web page. These surveys seem to be most useful as customer satisfaction surveys or marketing surveys. This type of systematic sampling can provide information that is generalizable to particular populations, such as those that visit a particular website/page. The surveys can be restricted to only those with certain IP (Internet Protocol) addresses, allowing one to target more specific subsets of visitors, and 'cookies' can be used to restrict the submission of multiple surveys from the same computer.

A potential issue with this type of survey is nonresponse. Coomly (2000) reports typical response rates in the 15 to 30 per cent range, with the lowest response rates occurring for poorly targeted and/or poorly designed surveys. The highest response rates were obtained for surveys that were relevant to the individual, either in terms of the particular survey questions or, in the case of marketing surveys, the commercial brand being surveyed.

As discussed in Couper (2000), an important issue with intercept surveys is that there is no way to assess nonresponse bias, simply because no information is available on those that choose not to complete a survey. Coomly (2000) hypothesizes that responses may be biased towards those who are more satisfied with a particular product, brand, or website; towards those potential respondents who are more computer and Internet savvy; and, away from heavy Internet users who are conditioned to ignore pop-ups. Another source of nonresponse bias for intercept surveys implemented as pop-up browser windows is pop-up blocker software, at least to the extent that pop-up blocker software is used differentially by various portions of the web-browsing community.

Pre-recruited panel surveys

Pre-recruited panel surveys are, generally speaking, groups of individuals who have agreed in advance to participate in a series of surveys. For online surveys requiring probability samples, these individuals are generally recruited via some means other than the web or e-mail – most often by telephone or postal mail. (See Toepoel, 2012, for guidance on how to build an online panel of respondents.)

For a longitudinal effort consisting of a series of surveys, researchers may recruit panel members

specifically for that effort. For smaller efforts or for single surveys, a number of companies maintain panels of individuals, pre-recruited via a probability-based sampling methodology, from which sub-samples can be drawn according to a researcher's specification. Knowledge Networks, for example, recruits all of its panel members via telephone using RDD, and it provides equipment and Internet access to those that do not have it in an attempt to maintain a panel that is a statistically valid cross section of the population (see Pineau and Dennis, 2004, and www.gfk.com/us/Solutions/consumer-panels/Pages/GfK-KnowledgePanel.aspx for additional detail). Other companies such as Qualtrics provide panel management software so that companies and researchers can assemble and manage their own panel or panels.[1]

Pre-recruited, Internet-enabled panels can provide the speed of online surveys while simultaneously eliminating the often-lengthy recruitment process normally required. As such, they can be an attractive option to researchers who desire to field an online survey, but who require a sample that can be generalized to populations outside of the Internet-user community.

However, pre-recruited panels are not without their potential drawbacks. In particular, researchers should be aware that long-term panel participants may respond differently to surveys and survey questions than first-time participants (called 'panel conditioning' or 'time-in-sample bias'). Also, nonresponse can be an issue if the combined loss of potential respondents throughout all of the recruitment and participation stages is significant. However, as Couper (2000) concludes, '… in theory at least, this approach begins with a probability sample of the full (telephone) population, and assuming no nonresponse error permits inference to the population…'.

Entertainment polls

Online entertainment polls are 'surveys' conducted purely for their entertainment value (though they are sometimes passed off to be more than what they are). On the Internet, they largely consist of websites where any visitor can respond to a posted survey. An example of an online entertainment poll is The Weekly Online Poll (www.arsenal.com/fansforum) for fans of the Arsenal football team. The telephone equivalent of these types of polls are call-in polls (or cellphone text message polls) such as those advertised on various television shows, where viewers can vote for their favourite contestant or character. Of course, online

---

[1] See, for example, www.qualtrics.com/panel-management/.

entertainment polls are as unscientific as call-in telephone polls.

Surveys using 'harvested' e-mail lists

Harvested e-mail lists are sets of e-mail addresses collected from postings on the web and from individuals who are (wittingly or unwittingly) solicited for their e-mail addresses. There are many commercial entities ('e-mail brokers') that sell lists of e-mail addresses or access to lists of e-mail addresses (just Google 'buy e-mail list'). Lists can also be assembled from resources on the web. In a similar vein, Yahoo e-mail addresses can be searched by name via the Yahoo Email Search (http://itools.com/tool/yahoo-email-address-search). Also similarly, Spokeo has an reverse e-mail search feature by email, username and address (www.spokeo.com/email-search).

Note that harvesting e-mail addresses and distributing unsolicited e-mail related to surveys could be a violation of professional ethical standards and/or illegal. For example, European Union Article 13(1) of the *Privacy and Electronic Communications Directive* prohibits the sending of unsolicited commercial e-mail. In a similar vein, the Council of American Survey Research Organizations (CASRO) Code of Standards and Ethics for Survey Research clearly states that using harvested e-mail addresses and sending unsolicited e-mail is unethical (www.casro.org/?page=TheCASROCode2014). For a more detailed discussion of the ethical considerations and implications, see Eynon et al., and Charlesworth (this volume), and Krishnamurthy (2002) and the references contained therein.

Samples derived from harvested e-mail lists are non-probability samples because they are based on a convenience sample of e-mail addresses. For example, in 2008 E-mail Marketing Blitz stated on their website that, 'Our targeted optin [sic] e-mail lists are updated monthly and gathered through special interest websites, entertainment websites and special alliances.' Such samples should not be confused with list-based probability samples where the e-mail addresses in the list-based sample represent a (virtually) complete list of the e-mail addresses of some target population.

The efficacy of sending unsolicited surveys to a list of purchased or otherwise procured e-mail addresses is questionable. Not only do e-mail addresses turn over quite frequently, but many of those on the list may have been recruited either without their knowledge, or they may have inadvertently agreed by failing to

uncheck a box when they signed up for something else. As a result, response rates are likely to be extremely low.

Unrestricted self-selected surveys

As with entertainment polls, unrestricted, self-selected surveys are surveys that are open to the public for anyone to participate in. They may simply be posted on a website so that anyone browsing through may choose to take the survey, or they may be promoted via website banners or other online advertisements, or they may be publicized in traditional print and broadcast media. Regardless of how they are promoted (or not), the key characteristics of these types of survey are that there are no restrictions on who can participate, and it is up to the individual to choose to participate (opt in).

For example, Berson et al. (2002) conducted a web-based survey 'to better understand the risks to adolescent girls online' by posting a link to their survey on the *Seventeen Magazine Online* website. Via the survey, the authors collected data on 10,800 respondents with 'identified behaviours that put them at risk'. The researchers were careful to appropriately qualify their results:

> The results highlighted in this paper are intended to explore the relevant issues and lay the groundwork for future research on youth in cyberspace. This is considered an exploratory study which introduces the issues and will need to be supplemented with ongoing research on specific characteristics of risk and prevention intervention. Furthermore, the generalizability of the study results to the larger population of adolescent girls needs to be considered. Due to anonymity of the respondents, one of the limitations of the research design is the possibility that the survey respondents did not represent the experience of all adolescent girls or that the responses were exaggerated or misrepresented.

Unrestricted, self-selected surveys are a form of convenience sampling and, as such, the results cannot be generalized to a larger population. But as Berson et al. illustrate, that does not necessarily negate their usefulness for research.

The web can also facilitate access to individuals who are difficult to reach either because they are hard to identify, locate, or perhaps exist in such small numbers that probability-based sampling would be unlikely to reach them in sufficient numbers. Coomber (1997) describes such a use of the web for fielding a survey to collect information from drug dealers about drug adulteration/dilution. By posting invitations to participate

in a survey on various drug-related discussion groups, Coomber collected data from 80 survey respondents (that he deemed reliable) located in 14 countries on four different continents. The sample was certainly not generalizable, but it also provided data that was unlikely to be collected in any other way, and which Coomber found consistent with other research.

In addition, Alvarez et al. (2002) proposed that these types of non-probability sample can be useful and appropriate for conducting experiments (say, in the design of web pages or web surveys) by randomly assigning members of the sample to control and experimental groups. In terms of psychology experiments, Siah (2005) states, 'For experimental research on the Internet, the advantage of yielding a heterogeneous sample seems persuasive considering that the most common criticism of psychological research is its over-reliance on college student samples.'

Opt-in panels (volunteer or paid)

Opt-in panels are similar in concept to the pre-recruited panels, except the volunteers are not recruited using a probability-based method. Rather, participants choose to participate, perhaps after coming across a solicitation on a website. In this regard, volunteer panels are similar to unrestricted, self-selected surveys except that those who opt in do so to take a continuing series of surveys. Harris Interactive manages such an opt-in panel. Its website (www.harrispollonline.com) banner says,, 'Influence Tomorrow – Get Awesome Rewards'.

Often these panels are focused on market research, soliciting consumer opinions about commercial products, and participants sometimes do it for monetary incentives. For example, the Harris website states,

> Once you become an HPOL member, you will instantly receive a quick survey to earn your first points! Once you complete the initial survey, you will start receiving invitations to participate in a variety of cool surveys. Topics include: consumer, technical, political, etc. For every survey you successfully complete, you will receive points called HIpoints, which you will be able to exchange for awesome rewards after just doing a couple of surveys! For each survey that you are directed to, but do not qualify for, you will be rewarded a HIstakes entry and 15 HIpoints. Also, with every survey you complete you have a chance to enter our $10,000 sweepstakes!

Researchers are also paying respondents to take surveys on sites like Mechanical Turk

(www.mturk.com/mturk/welcome). On these sites, surveys (and other tasks) are posted and people complete the tasks for a price. As reported on NPR's All Tech Considered (2014),

> You can buy just about anything on Amazon.com — clothes, books, electronics. You can buy answers, too. College students and professors are doing all sorts of research on an Amazon site called Mechanical Turk.

> Need 200 smokers for your survey on lung cancer? Have a moral dilemma to pose for your paper on Kierkegaard? Now researchers can log in, offer a few pennies in payment and watch the data roll in.

The article goes on to say,

> Berinsky, over at MIT, says researchers save not just weeks of effort, but great amounts of money as well. He's pays a couple cents per participant, compared to the $10 or $15 he used to pay. OK, so MTurk is fast and cheap. Is it good? How reliable are the data? The research shows that the population of Turkers is pretty representative, more so than signing up college students.

However, as "Steven O" says in the comments section of the website,

> Selection bias is still an issue - sample size doesn't help if you have a non-random sample. For alpha testing, fine, but I would be very interested to see a study showing how Mechanical Turk respondents (those online and motivated to take a survey by at most $2.50/hour) represent most target populations. "Better than college students" is a low bar to beat.

The issue, as with all non-probability samples, is that those taking the survey are unlikely to be representative of the population of inference. For example, those working on the Mechanical Turk site are likely to be younger, more computer literate, and of a lower socioeconomic strata than the general population. Furthermore, on sites like this, it is difficult to limit the respondents to a specific geographic location or nationality. Whether this is a problem depends on the objectives of the specific research effort.

ISSUES AND CHALLENGES IN INTERNET-BASED SURVEY SAMPLING

All survey modes have their strengths and weaknesses; online surveys are no different in this regard. The various strengths and weaknesses are more or less important, depending on the survey's purpose. Drawing an appropriate sample that will provide the data necessary to appropriately address the research objective is

critical. Hence, in this section we focus on the issues and challenges related to sampling for online surveys.

Sampling frame and coverage challenges

A frequent impediment for conducting large-scale, online surveys is the lack of a sampling frame. Simply put, no single registry or list of e-mail addresses exists and thus list-based sampling frames are generally available only for specific populations (government organizations, corporations, etc).

Compounding this difficulty, and leaving aside the issue of population coverage to be discussed shortly, it is impossible to employ a frameless sampling strategy, since for all practical purposes one cannot assemble random e-mail addresses. Of course, it is theoretically possible to 'construct' e-mail addresses by repeatedly randomly concatenating letters, numbers, and symbols, but the sheer variety of e-mail addresses means most of the constructed addresses will not work. More importantly, the unstructured nature of the Internet means that even if one could tolerate the multitude of undeliverable e-mail messages that would result, they would not be useful as the basis for a probability sample.

In terms of coverage, it is widely recognized that online surveys using only samples of Internet users do not generalize to the general public. While Internet penetration into households continues at a rapid pace (Figure 14.3 shows the top 20 countries with the highest Internet penetration as of the end of 2013), the penetration is far from complete (compared to, say, the telephone) and varies widely by country and region of the world.[3] The point is, if the target of inference is the general public, considerable coverage error remains for any sample drawn strictly from Internet users. Furthermore, it is worth emphasizing that even with 100 per cent Internet penetration, the difficulty of obtaining a true probability sample from the general population remains.

#	Country or Region	Population, 2014 Est	Internet Users Year 2000	Internet Users Dec 2013	Penetration (% Population)	% Growth 2000 - 2013
	TOP 20 COUNTRIES WITH HIGHEST NUMBER OF INTERNET USERS - December 31, 2013					
1	China	1,355,692,576	22,500,000	620,907,200	45.8 %	2,659.6 %
2	United States	318,892,103	95,354,000	268,507,150	84.2 %	181.6 %
3	India	1,236,344,631	5,000,000	195,248,950	15.8 %	3,805.0 %
4	Brazil	202,656,788	5,000,000	109,773,650	54.2 %	2,095.5 %
5	Japan	127,103,388	47,080,000	109,626,672	86.2 %	132.9 %
6	Russia	142,470,272	3,100,000	87,476,747	61.4 %	2,721.8 %
7	Germany	80,996,685	24,000,000	69,779,160	86.2 %	190.7 %
8	Nigeria	177,155,754	200,000	67,319,186	38.0 %	33,559.6 %
9	United Kingdom	63,742,977	15,400,000	57,266,690	89.8 %	271.9 %
10	France	66,259,012	8,500,000	55,221,000	83.3 %	549.7 %
11	Indonesia	253,609,643	2,000,000	55,000,000	21.7 %	2,650.0 %
12	Mexico	120,286,655	2,712,400	52,276,580	43.5 %	1,827.3 %
13	Iran	80,840,713	250,000	45,000,000	55.7 %	17,900.0 %
14	Philippines	107,668,231	2,000,000	44,200,540	41.1 %	2,110.0 %
15	Egypt	86,895,099	450,000	43,065,211	49.6 %	9,470.0 %
16	Korea	49,039,986	19,040,000	41,571,196	84.8 %	118.3 %
17	Vietnam	93,421,835	200,000	41,012,186	43.9 %	20,406.1 %
18	Turkey	81,619,392	2,000,000	37,748,969	46.3 %	1,787.4 %
19	Italy	61,680,122	13,200,000	36,058,199	58.5 %	173.2 %
20	Spain	47,737,941	5,387,800	35,705,960	74.8 %	562.7 %
	TOP 20 Countries	4,754,113,803	273,374,200	2,072,765,246	43.6 %	658.2 %
	Rest of the World	2,427,744,816	87,611,292	729,713,688	30.1 %	732.9 %
	Total World Users	7,181,858,619	360,985,492	2,802,478,934	39.0 %	676.3 %

Figure 14.3 Top 20 countries with the highest Internet penetration as of December 31, 2013. (Source: Internet World Stats, 2015a)

Now, even if there is minimal coverage error for a particular online survey effort, when using only an online survey mode the target population must also be sufficiently computer literate and have both regular and easy access to the Internet to facilitate responding to the survey. Simply put, just because an organization maintains a list of e-mail addresses for everyone in the organization it does not necessarily follow that every individual on the list has equal access. Lack of equal access could result in significant selection and nonresponse biases.

Mixed-mode surveys using online and traditional media

For some surveys it may be fiscally and operationally possible to contact respondents by some mode other than e-mail, such as mail or telephone. In these cases the survey target population can be broader than that

for which an e-mail sampling frame is available, up to and including the general population. But at present such a survey must also use multiple survey modes to allow respondents without Internet access the ability to participate. Mixed-mode surveys may also be useful for alleviating selection bias for populations with uneven or unequal Internet access, and the sequential use of survey modes can increase response rates.

For example, Dillman (2007: 456) describes a study in which surveys that were fielded using one mode were then followed up with an alternate mode three weeks later. As shown in Table 14.3, in all cases the response rate increased after the follow-up. Now, of course, some of this increase can be attributed simply to the fact that a follow-up effort was conducted. However, the magnitude of the increases also suggests that offering a different response mode in the follow-up can be beneficial.

However, mixed-mode surveys are subject to other issues. Two of the most important are mode effects and respondent mode preferences. Mode effects arise when the type of survey affects how respondents answer questions. Comparisons between online surveys and traditional surveys have found conflicting results, with some researchers reporting mode effects and others not. See, for example, the discussion and results in Schonlau et al. (2004: 130). Though not strictly a sampling issue, the point is that researchers should be prepared for the existence of mode effects in a mixed-mode survey. Vehovar and Manfreda's overview chapter (this volume) explores in greater detail the issues of combining data from online and traditional surveys.

In addition, when online surveys are part of a mixed-mode approach, it is important to be aware that the literature currently seems to show that respondents will tend to favour the traditional survey mode over an online mode. See, for example, the discussions in Schonlau et al. (2002) and Couper (2000: 486–487). Fricker and Schonlau (2002), in a study of the literature on web-based surveys, found 'that for most of the studies respondents currently tend to choose mail when given a choice between web and mail. In fact, even when respondents are contacted electronically it is not axiomatic that they will prefer to respond electronically'.

The tendency to favour non-online survey modes lead Schonlau et al. (2002: 75) to recommend for mixed-mode mail and web surveys that:

… the most effective use of the Web at the moment seems to involve a sequential fielding scheme in which respondents are first encouraged to complete the survey via the Web and then nonrespondents are subsequently sent a paper survey in the mail. This approach has the advantage of maximizing the potential for cost savings from using Internet while maintaining the population coverage and response rates of a mail survey.

Table 14.3 As reported in Dillman (2007), using an alternate survey mode as a follow-up to an initial survey mode can result in higher overall response rates

Initial survey mode and response rate	Follow-up survey mode and combined response rate	Response rate increase
Mail (75%)	Telephone (83%)	8%
Telephone (43%)	Mail (80%)	37%
IVR[4] (28%)	Telephone (50%)	22%
Web (13%)	Telephone (48%)	35%

Web-based recruitment issues and effects

Whether e-mail addresses are constructed, assembled from third-party sources, or harvested directly from the web, there is the issue of unsolicited survey e-mail as spam. For example, Sheehan (1999) conducted a survey with e-mail addresses harvested from www.Four11.com and stated, 'Several individuals receiving the solicitation e-mail censured the researchers for sending out unsolicited e-mails, and accused the researchers of "spamming". ' They further recounted that 'One [ISP] system operator [who observed a large number of e-mail messages originating from a single address] then contacted his counterpart at our university.'

In addition, distributing an unsolicited online survey is also not without its perils. For example, Andrews et al. (2002) report on a study of 'hard-to-involve Internet users' – those who lurk in, but do not participate publicly in, online discussion forums. In their study, an invitation to participate in a web survey was posted as a message to 375 online community discussion boards. While they collected 1,188 valid responses (out of 77,582 discussion board members), they also 'received unsolicited e-mail offers, some of which were pornographic in content or aggressive in tone' and they had their web server hacked twice, once with the infection of a virus.

In spite of the challenges and possible perils, it is possible to recruit survey participants from the web. For

example, Alvarez et al. (2002) conducted two online recruitment efforts – one using banner advertisements on web pages and another using a subscription check box. In brief, their results were as follows.

- In the first recruitment effort, Alvarez et al. ran four 'banner' campaigns in 2000 with the intention of recruiting survey participants using web-page banner advertisements. In the first campaign, which is representative of the other three, an animated banner advertisement resulted in more than 3.5 million 'impressions' (the number of times the banner was displayed), which resulted in the banner being clicked 10,652 times, or a rate of 3 clicks per 1,000 displays. From these 10,652 clicks, 599 survey participants were recruited.

- In the second recruitment effort, the authors ran a 'subscription' campaign in 2001 in which they arranged with a commercial organization to have a check box added to subscription forms on various websites. Essentially, Internet users who were registering for some service were given an opportunity to check a box on the service's subscription form indicating their willingness to participate in a survey. As part of this effort, the authors conducted two recruitment drives, each of which was intended to net 10,000 subscriptions. Across the two campaigns, 6,789 new survey participants were obtained from 21,378 subscribers.

The good news from the Alvarez et al. (2002) study is that, even though the banner approach yielded fewer new survey participants, both methods resulted in a significant number of potential survey respondents over a relatively short period of time: 3,431 new subjects over the course of six or seven weeks from the banner campaigns, and 6,789 new subjects over the course of three weeks from the subscription campaigns. Each banner subject cost about \$7.29 to recruit, while the subscription subjects cost only \$1.27 per subject. (Unfortunately, the authors did not present any data on survey completion rates, so we do not know whether there were differences between the two samples that might have favored one over the other).

The bad news is that the two groups differed significantly in all of the demographic categories collected (gender, age, race, and education) and they differed in how they answered questions on exactly the same survey. In addition, both groups differed significantly from the demographics of the Internet population as measured by the August 2000 Current Population Survey. The problem, of course, is that there are clear effects associated with how subjects are recruited, such that the resulting samples are different even from the general Internet population. Shillewaert et al. (1998) found similar recruitment method biases. Hence, while it is possible to ethically recruit survey participants from the web, it seems that the recruitment methodology

affects the types of individual that self-select into the sample.

Improving response rates for online surveys

Response rates have a direct effect on sampling: the higher the response rate, the fewer people need to be sampled to achieve a desired number of survey completions. In addition, higher response rates are associated with lower nonresponse bias.

Unfortunately, in a summary of the academic survey-related literature up through 2001, Fricker and Schonlau (2002) concluded that 'Web-only research surveys have currently only achieved fairly modest response rates, at least as documented in the literature.' S. Fricker et al. (2005) similarly summarized the state of affairs as 'Web surveys generally report fairly low response rates.'

A good illustration of this is the Couper et al. (1999) study in which employees of five US federal government statistical agencies were randomly given a mail or e-mail survey. Comparable procedures were used for both modes, yet higher response rates were obtained for mail (68–76 per cent) than for e-mail (37–63 per cent) across all of the agencies.

Incentives are a common and effective means for increasing response rates in traditional surveys. Goritz (2006) is an excellent review of the use of incentives in survey research in which he distinguishes their use in traditional surveys from online surveys and provides a nice discussion of the issues associated with using incentives in online surveys. Open issues include:

- how best to deliver an incentive electronically;

- whether it is better to provide the incentive prior to a respondent taking the survey or after;

- whether incentives have different effects for individuals taking a survey one time versus pre-recruited panel members who take a series of surveys.

Individual studies of online surveys have generally found incentives to have little or no effect. For example, Coomly (2000) found that incentives had little effect on response rates for pop-up surveys, and Kypri and Gallagher (2003) found no effect in a web-based survey. However, Göritz (2006) conducted a meta-analysis of 32 experiments evaluating the impact of incentives on survey 'response' (the fraction of those solicited to

take the survey that actually called up the first page of the survey), and 26 experiments evaluating the effect of incentives on survey 'retention' (the fraction of those who viewed the first page that actually completed the survey). From the meta-analysis, Görtiz concluded that 'material incentives promote response and retention in Web surveys' where 'material incentives increase the odds of a person responding by 19% over the odds without incentives' and 'an incentive increased retention by 4.2% on average'.

In addition to incentives, Dillman (2007) and Dillman et al. (1999) have put forward a number of survey procedural recommendations to increase survey response rates, based on equivalent methods for traditional surveys, which we will not re-cover here since they are mainly related to survey design and fielding procedures. While we do note that the recommendations seem sensible, Couper (2000) cautions that 'there is at present little experimental literature on what works and what does not'.

Bigger samples are not always better

With online surveys using a list-based sampling frame, rather than sending the survey out to a sample, researchers often simply send the survey out to the entire sampling frame. That is, researchers naively conducting (all electronic) online surveys – where the marginal costs for additional surveys can be virtually nil – often fail to recognize the 'trade-off between easy, low cost access to large numbers of patients [participants] and the representativeness in the population being studied' (Soetikno et al., 1997). As we previously discussed, for both probability and non-probability-based samples, larger sample sizes do not necessarily mean the sample is more representative of any greater population: a sample can be biased whether it is large or small.

One might argue that in these situations the researchers are attempting to conduct a census, but in practice they are forgoing a probability sample in favour of a convenience sample by allowing members of the sampling frame to opt into the survey. Dillman et al. (1999) summarized this practice as follows: '…the ease of collecting hundreds, thousands, or even tens of thousands of responses to web questionnaires at virtually no cost, except for constructing and posting, appears to be encouraging a singular emphasis on the reduction of sampling error'. By this Dillman et al. mean that researchers who focus only on reducing sampling error by trying to collect as large a sample as possible miss the point that it is equally important to reduce coverage, measurement, and nonresponse error in order to be able to accurately generalize from the sample

data.

A myopic focus on large sample sizes – and the idea that large samples equate to sample representativeness which equates to generalizability – occurs with convenience sample-based web and e-mail surveys as well. 'Survey2000' is an excellent example of this type of focus. A large-scale, unrestricted, self-selected survey, conducted as a collaborative effort between the National Geographic Society (NGS) and some academic researchers, Survey2000 was fielded in 1998. The survey was posted on the National Geographic Society's website and participants were solicited both with a link on the NGS homepage and via advertisements in NGS periodicals, other magazines, and newspapers.

Upon completion of the effort, Witte et al. (2000) report that more than 80,000 surveys were initiated and slightly more than 50,000 were completed. While this is an impressively large number of survey completions, the unrestricted, self-selected sampling strategy clearly results in a convenience sample that is not generalizable to any larger population. Yet, Witte et al. (2000) go to extraordinary lengths to rationalize that their results are somehow generalizable, while simultaneously demonstrating that the results of the survey generally do not correspond to known population quantities.

Misrepresenting convenience samples

A related and significant concern with non-probability-based sampling methods, both for online and traditional surveys, is that survey accuracy is characterized only in terms of sampling error and without regard to the potential biases that may be present in the results. While this has always been a concern with all types of survey, the ease and spread of online surveys seems to have exacerbated the practice. For example, the results of an 'E-Poll' were explained as follows:

> THE OTHER HALF / E-Poll® Survey of 1,007 respondents was conducted January 16–20, 2003. A representative group of adults 18+ were randomly selected from the E-Poll online panel. At a 95% confidence level, a sample error of +/− 3% is assumed for statistics based on the total sample of 1,007 respondents. Statistics based on sub-samples of the respondents are more sensitive to sampling error. (From a press release posted on the E-Poll website.)

No mention was made in the press release that the 'E-Poll online panel' consists of individuals who had chosen to participate in online polls, nor that they were unlikely to be representative of the general

population. Rather, it leaves readers with an incorrect impression that the results apply to the general population when, in fact, the margin of error for this particular survey is valid only for adult members of the E-Poll online panel.

In response to the proliferation of such misleading statements, the American Association for Public Opinion Research (AAPOR) has publicly stated that 'The reporting of a margin of sampling error associated with an opt-in or self-identified sample (that is, in a survey or poll where respondents are self-selecting) is misleading.' They go on to say, 'AAPOR considers it harmful to include statements about the theoretical calculation of sampling error in descriptions of such studies, especially when those statements mislead the reader into thinking that the survey is based on a probability sample of the full target population' (AAPOR, 2007).

IN SUMMARY

Every survey effort can be classified according to how the respondents are contacted (the contact mode), how they are asked to complete the survey (the response mode), and then how subsequent communication is conducted (the follow-up mode). Each of these can be executed in different media, where the media are telephone, mail, web, e-mail, and so forth. For example, respondents may be contacted by telephone to participate in a web survey with follow-up done by mail.

Explicitly specifying contact, response, and follow-up modes is often irrelevant for traditional surveys, since respondents that have been asked to take, say, a telephone survey have generally been contacted via the same mode. While not a strict rule – for example, a telephone survey may be preceded by mailed invitation to each survey respondent – it is often the case. In comparison, given the challenges that we have discussed in this chapter, the contact, response, and follow-up modes are much more likely to differ with online surveys.

In terms of sampling for online surveys, what is relevant is that the sampling methodology is generally driven by the contact mode, *not* the response mode. Hence, as shown in Table 14.4, we can organize sampling strategies by contact mode, where the check marks indicate which sampling strategies are mainly associated with the various contact methods.

Note that we are focusing explicitly on an online survey response mode in Table 14.4. So, for example, while systematic sampling can be applied to phone or mail surveys, the telephone is not likely to be used as a contact medium for an online survey using systematic sampling, and hence those cells in the table are not checked. Similarly, while there is a plethora of phone-in entertainment polls, neither the telephone nor postal mail is used to contact respondents to take online entertainment polls.

From Table 14.4 we can broadly summarize the current state of the art for the various online survey methods and their limitations as follows.

- Entirely web-based surveys, meaning surveys in which the potential respondents are contacted on the web and take a web survey, are chiefly limited to collecting data from non-probability-based samples.

  o The exception is systematic sampling for pop-up/intercept surveys that are predominantly used for customer-satisfaction types of survey associated with specific websites or web pages.

  o Respondent contact for online surveys using non-probability samples can also be conducted via traditional (non-online) media and advertising.

- Research surveys that require probability sampling are very limited when using an online contact mode (web and e-mail).

  o E-mail is useful as a contact mode only if a list of e-mail addresses is available. Such a list is an actual or de facto sampling frame, from which a sample may be drawn or a census attempted.

  o The population of inference is usually quite limited when using an e-mail address sampling frame. It is generally the sampling frame itself.

  o A poorly conducted census of an entire e-mail list may limit the survey results even further, since nonresponse and other biases may preclude generalizing even to the sample frame.

- If the research objectives require inferring from the survey results to some general population, then respondents will most likely have to be contacted by a non-online medium.

  o If the population of inference is a population in which some of the members do not have e-mail/web access, then the contact mode will have to be a non-online medium.

  o Under such conditions, the survey will have to be conducted using a mixed mode, so that those without Internet access can participate. Conversely, lack of a non-online survey mode will result

- o Pre-recruited panels can provide ready access to pools of online survey respondents, but to allow generalization to some larger, general population such panels need to be recruited using probability sampling methods from the general population (usually via RDD). And, even under such conditions, researchers need to carefully consider whether the panel is likely to be subject to other types of bias.

Table 14.4 Sampling strategies for online surveys by contact mode

			Sampling Strategy								
			Probability-based					**Non-probability-based**			
			List-based sampling frames	*Non-list-based random sampling*	*Systematic sampling*	*Mixed-mode survey with online option*	*Pre-recruited survey panel*	*Entertainment polls*	*Unrestricted self-selected surveys*	*Harvested e-mail lists*	*Volunteer (opt-in) panels*
Contact Method	**Online**	Web			✓			✓	✓		✓
		E-mail	✓							✓	
	Offline	Telephone	✓	✓		✓	✓				
		Postal Mail	✓			✓					
		In-person				✓					
		Other: TV, print advertising , etc.						✓	✓		✓

AAPOR online panel recommendations

In 2008, the American Association of Public Opinion Research (AAPOR) Executive Council established an Opt-in Online Panel Task Force with the charge of "reviewing the current empirical findings related to opt-in online panels utilized for data collection and developing recommendations for AAPOR members" (Baker et al., 2010). Among others, the report made the following recommendations:

- "Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values.

- Although mode effects may account for some of the differences observed in comparative studies, the use of nonprobability sampling in surveys with online panels is likely the more significant factor in the overall accuracy of surveys using this method.

- There are times when a nonprobability online panel is an appropriate choice.

- Research aimed at evaluating and testing techniques used in other disciplines to make population inferences from nonprobability samples is interesting but inconclusive" (Baker et al, 2012: 758-759).

Looking to the future

So, what does the future hold for online survey sampling? At this point in the Internet's development, with its rapid expansion and continued evolution, it's truly impossible to say. Yet we can hazard a few guesses.

First, if the Internet continues to expand but largely maintains its current structure, then advances in sampling methods that will allow random sampling of, and inference to, general populations will be at best slow and difficult to develop. This follows from the fact that Internet-wide sampling frames are simply unavailable under the current Internet structure/organization, and no general frameless sampling strategies yet exist. Unless the way the Internet is organized and operated changes, it seems this will continue to be the case into the foreseeable future.

That said, survey methodologists should endeavour to develop new sampling paradigms for online surveys. The fundamental requirement for a probability-based sampling scheme is that every member of the target population has a known, non-zero probability of being sampled. While in traditional surveys this can be achieved via various frame and frameless sampling strategies, it does not necessarily follow that online surveys must use those same sampling strategies. Rather, new sampling methods that take advantage of the unique characteristics of the Internet, such as the near-zero marginal cost for contacting potential respondents, should be explored and developed.

In addition, researchers considering conducting an online survey should consider whether the capabilities of the web can be leveraged to collect the desired data in some other innovative fashion. For example,

Lockett and Blackman (2004) present a case study of Xenon Laboratories, an Internet-based financial services firm that employed a novel approach to market research. Xenon Laboratories wanted to collect data on foreign exchange charges by credit card companies on business travellers. They recognized that neither the travellers nor the credit card companies were likely to respond to a survey on this topic, whether fielded over the web or otherwise. Instead Xenon Laboratories developed the Travel Expenses Calculator (www.xe.com/tec) and the Credit Card Charges Calculator (www.xe.com/ccc) and posted them on the web for anyone to use for free. These tools help foreign business travellers to accurately calculate the cost of a business expense receipt in terms of their own currency. Lockett and Blackman (2004) say, 'On the basis of this information [input by those using the calculators] it is possible to conduct basic market research by aggregating the inputted calculations. Xenon is now in the unique position to analyse whether or not the different card providers employ the same charging levels and whether or not these companies' charge structures vary according to geographical region.' They go on to conclude, 'This value-added approach, which is mutually beneficial to both parties, is an important and novel approach to market research.'

Second, it is also possible that technological innovation will facilitate other means of sampling for online surveys (and for conducting the surveys themselves, for that matter). For example, current trends seem to point towards a merging of the Internet with traditional technologies such as television and telephone. Indeed, all of these services are merging into one common household device through which a consumer can simultaneously watch television, surf the web, send e-mail, and place telephone calls – otherwise known as the smartphone. Depending on how smartphone and related technology evolves, various types of random sampling methodologies, as well as new survey modes, may become possible. For example, it may become feasible, and perhaps even desirable, to sample respondents via RDD and then send the potential respondent a text message with an embedded URL to a web survey. Or, perhaps via RDD a survey interviewer calls the potential respondent and then in real time the respondent completes an interviewer-assisted, web-based survey. Or, sometime in the future, it may also be possible to use smartphone companies' subscriber listings as sampling frames (much as telephone directories were used pre-RDD in the mid-1900s). Or it may be that some other state emerges that lends itself to some form of sampling that is not possible today. The point is that the Internet is still very much in its infancy, and the current difficulties surrounding sampling for online

surveys described in this chapter may or may not continue into the future.

To put this in a historical context, note that while the telephone was invented in the late 1800s, and telephone systems developed and expanded rapidly through the early 1900s, it was not until the mid-1900s that telephone coverage was sufficiently large and standards for telephone numbers adopted that made RDD possible. In fact, the foundational ideas for an efficient RDD sampling methodology were not proposed until the early 1960s (Cooper, 1964), after which it took roughly another decade of discussion and development before RDD as we know it today became commonplace.[5] Thus, in total, it was roughly a *century* after the invention of the telephone before RDD became an accepted sampling methodology.

In comparison, the web has been in existence, in a commercial sense, only for little more than a decade or two. As with the telephone in the late 1800s and early 1900s, we are in a period of technological innovation and expansion with the Internet. However, unlike the telephone, given today's pace of innovation, the Internet and how we use it is likely to be quite different even just a few years from now. How this affects sampling for online surveys remains to be seen.

NOTES

1 Readers interested in the mathematics should consult one of the classic texts such as Kish (1965) or Cochran (1977); readers interested in a summary treatment of the mathematics and/or a more detailed discussion of the sampling process may consult a number of other texts, such as Fink (2003) or Fowler (2002). For those specifically interested in sampling methods for qualitative research, see Patton (2002).

2 These were not the only errors made in the 1936 and 1948 US presidential election polls (for more detail, see Zetterberg (2004) or Ross (1977)), but they were significant errors, and are highlighted here to illustrate that the various biases that can creep into survey and data collection can be both subtle and non-trivial.

3 For example, as of December 31, 2013 Internet World Stats (2015b) reported that the top 50 countries and regions of the world had a combined Internet penetration of 84.4 per cent, ranging from a high of 96.9 per cent for the Falkland Islands to 73.9 per cent for the Puerto Rico. In comparison, Internet penetration for the rest of the world was estimated to be 31.6 per cent.

4 IVR stands for Interactive Voice Response. These are automated telephone surveys in which prerecorded questions are used and respondents' answers are collected using voice-recognition technology.

5 See, for example, 'Random Digit Dialing as a Method of Telephone Sampling' (Glasser and Metzger, 1972), 'An Empirical Assessment of Two Telephone Sampling Designs' (Groves, 1978), and 'Random Digit Dialing: A Sampling Technique for Telephone Surveys' (Cummings, 1979).

REFERENCES

Alvarez, R.M., Sherman, R.P. and VanBeselaere, C. (2002) 'Subject acquisition for web-based surveys', dated September 12, 2002, accessed online at *survey.caltech.edu/alvarez.pdf* on September 29, 2006.

American Association for Public Opinion Research (AAPOR) (2007) 'Reporting of margin of error or sampling error in online and other surveys of self-selected individuals'. Accessed online at www.aapor.org/pdfs/2006/samp_err_stmt.pdf on April 23, 2007.

Andrews, D., Nonnecke, B. and Preece, J. (2002) 'Electronic survey methodology: a case study in reaching hard-to-involve Internet users', *International Journal of Human-Computer Interaction*, 12 (2): 185–210.

Baker, R., Blumberg, S.J., Brick, J.M., Couper, M.P., Courtright, M., Dennis, J.M., Dillman, D., Frankel, M.R., Garland, P., Groves, R.m., Kennedy, C., Krosnick, J., Lavrakas, P.J., Lee, S., Link, M., Piekarski, L., Rao, K., Thomas, R.K., and D. Zahs (2010). Reserch Synthesis: AAPOR Report on Online Panels, *Public Opinion Quarterly*, **74**, 711—781.

Berson, I.R., Berson, M.J. and Ferron, J.M. (2002) 'Emerging risks of violence in the digital age: lessons for educators from an online study of adolescent girls in the United States', *Journal of School Violence*, 1 (2), 51–72. Published simultaneously online in Meridian, A Middle School Computer Technologies Journal, accessed at www2.ncsu.edu/unity/lockers/project/meridian/sum2002/cyberviolence/cyberviolence.pdf on September 29, 2006.

Callegaro, M., Baker, R., Bethlehem, J., Goritz, A.S., Krosnick, J.A., and P.J. Lavrakas (2014). *Online Panel Research: A Data Quality Perspective*, Wiley Series in Survey Methodology, John Wiley & Sons.

Cochran, William G. (1977) *Sampling Techniques*. New York: John Wiley.

Coomley, P. (2000) 'Pop-up surveys: what works, what doesn't work and what will work in the future.' Proceedings of the ESOMAR worldwide Internet conference Net Effects 3, volume 237. Amsterdam, NL: ESOMAR. Accessed at www.virtualsurveys.com/ news/papers/paper_4.asp on September 21, 2006.

Coomber, R. (1997) 'Using the Internet for Survey Research', *Sociological Research Online*, 2, 14–23.

Cooper, S.L. (1964) 'Random sampling by telephone: an improved method', *Journal of Marketing Research*, 1 (4): 45–8.

Couper, Mick P. (2000) 'Review: web surveys: a review of issues and approaches', *The Public Opinion Quarterly*, 64 (4): 464–94.

Couper, Mick P., Blair, J. and Triplett, T. (1999) 'A comparison of mail and e-mail for a survey of employees in federal statistical agencies', *Journal of Official Statistics*, 15 (1): 39–56.

Cummings, K.M. (1979) 'Random Digit Dialing: a sampling technique for telephone surveys', *Public Opinion Quarterly*, 43, 233–44.

Dillman, D.A. (2007) *Mail and Internet Surveys: The Tailored Design Method* (2007 update with new Internet, visual, and mixed-mode guide), 2nd edn. New York: John Wiley.

Dillman, D.A., Tortora, R.D. and Bowker, D. (1999) 'Principles for constructing web surveys', accessed online at www.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf on January 27, 2005.

Fink, Arlene (2003) *How to Sample in Surveys*, 2nd edn: The Survey Kit, volume 7. Thousand Oaks: Sage.

Fowler, Jr. and Floyd J. (2002) *Survey Research Methods*, 3rd edn: Applied Social Research Methods Series, volume 1. Thousand Oaks: Sage.

Fricker, Jr., R.D. and Schonlau, M. (2002) 'Advantages and disadvantages of Internet research surveys: evidence from the literature', *Field Methods*, 14: 347–67.

Fricker, S., Galesic, M., Tourangeau, R. and Yan, T. (2005) 'An experimental comparison of web and telephone surveys', *Public Opinion Quarterly*, 69 (3): 370–92.

Glasser, G.J. and Metzger, G.D. (1972) 'Random-Digit Dialing as a method of telephone sampling', *Journal of Marketing Research*, 9 (1), 59–64.

Göritz, A.S. (2006) 'Incentives in web studies: methodological issues and a review', *International Journal of Internet Science*, 1 (1): 58–70.

Groves, R.M. (1978) 'An empirical comparison of two telephone sample designs', *Journal of Marketing Research*, 15 (4), 622–31.

Groves, R.M. (1989) *Survey Errors and Survey Costs*. New York: John Wiley.

Internet World Stats (2015b) Accessed online at www.internetworldstats.com/top20.htm on May 10, 2015.

Internet World Stats (2015b) Accessed online at www.internetworldstats.com/top25.htm on May 10, 2015.

Jackob, N., Arens, J. and Zerback, T. (2005) 'Sampling procedure, questionnaire design, online implementation, and survey response in a multinational online journalist survey'. Paper presented at the Joint WAPOR/ISSC Conference: Conducting International Social Surveys. Accessed online at ciss.ris.org/uploadi/editor/1132070316WAPORPaper.pdf on September 30, 2006.

Kish, Leslie (1965) *Survey Sampling*. New York: Wiley-Interscience. (New edition February 1995).

Krishnamurthy, S. (2002) 'The ethics of conducting e-mail surveys'. Readings in virtual research ethics: issues and controversies, accessed online at ssrn.com/ abstract_id=651841 on September 30, 2006.

Kypri, K. and Gallagher, S.J. (2003) 'Incentives to increase participation in an Internet survey of alcohol use: a controlled experiment', *Alcohol and Alcoholism*, 38 (5): 437–441.

Lockett, A. and Blackman, I. (2004) 'Conducting market research using the Internet: the case of Xenon Laboratories', *Journal of Business and Industrial Marketing*, 19 (3), 178–87.

Library of Congress (2012). Figure contained in "Stop the Presses!" an online blog by Erin Allen dated November 6, 2012. Accessed online at http://blogs.loc.gov/loc/2012/11/stop-the-presses/ on May 16, 2015.

Midwest Book Review (2003) 'Book review: Take advantage of the Internet and preserve data integrity', accessed online at www.amazon.com/Conducting-Research-Surveys-E-Mail-Web/dp/0833031104/sr=1-3/qid=1160191573/ref=sr_1_3/102-5971652-1424906?ie=UTF8&s=books on October 6, 2006.

NPR (2014) 'Post A Survey On Mechanical Turk And Watch The Results Roll In', accessed online at www.npr.org/blogs/alltechconsidered/2014/03/05/279669610/post-a-survey-on-mechanical-turk-and-watch-the-results-roll-in

Patton, M.Q. (2002) *Qualitative Evaluation and Research Methods*, London: Sage.

Pineau, V. and Dennis, J.M. (2004) 'Methodology for probability-based recruitment for a web-enabled panel', dated November 21, 2004. Downloaded from www.knowledgenetworks.com/ganp/reviewer-info.html on July 21, 2006.

Ross, I. (1977) *The Loneliest Campaign: The Truman Victory*, (reprint edition). Greenwood Press. Accessed online at www.trumanlibrary.org/whistlestop/study_collections/1948campaign/large/docs/documents/index.php?documentdate=1968-00-00&documentid=53&studycollectionid=Election&pagenumber=1 on October 1, 2006.

Schonlau, Matthias, Fricker, Jr. Ronald D. and Elliott, Marc N. (2002) *Conducting Research Surveys via E-Mail and the Web*, MR-1480-RC, Santa Monica: RAND.

Schonlau, Matthias, Zapert, K., Simon, L.P., Sanstad, K.H., Marcus, S.M., Adams, J., Spranca, M., Kan, H., Turner, R. and Berry, S.H. (2004) 'A comparison between responses from a propensity-weighted web survey and an identical RDD survey', *Social Science Computer Review*, 22 (1): 128–38.

Sheehan, K.B. (1999) 'Using e-mail to survey Internet users in the United States: methodology and assessment', *Journal of Computer Mediated Communication*, (4) 3. Accessed online at http://jcmc.indiana.edu/vol14/issue3/sheehan.html on July 6, 2006.

Shillewaert, N., Langerak, F. and Duhamel, T. (1998) 'Non-probability sampling for WWW surveys: a comparison of methods', *Journal of the Market Research Society*, 40 (4): 307–22.

Siah, C.Y. (2005) 'All that glitters is not gold: examining the perils and obstacles in collecting data on the Internet', *International Negotiation*, 10: 115–30.

Soetikno, R.M., Provenzale, D. and Lenert, L.A. (1997) 'Studying ulcerative colitis over the World Wide Web', *American Journal of Gastroenterology*, 92 (3): 457–60.

Squires, P. (1988) 'Why the 1936 "Literary Digest" poll failed', *The Public Opinion Quarterly*, 53 (1): 125–33.

Toepoel, V. (2012). Building Your Own online Panel Via E-mail and Other Digital Media in *Handbook of Survey Methodology for the Social Sciences*, L. Gideon, ed., Springer.

Witte, J.C., Amoroso, L.M. and Howard, P.E.N. (2000) 'Research methodology: method and representation in Internet-based survey tools – mobility, community, and cultural identity in Survey2000', *Social Sciences Computer Review*, 18 (2): 179–95.

Wright, T. and Tsao, H.J. (1983) 'A frame on frames: an annotated bibliography'. In T. Wright (ed.) *Statistical Methods and the Improvement of Data Quality*. New York: Academic Press.

Zetterberg, H.L. (2004) 'US election 1948: the first great controversy about polls, media, and social science'. Paper presented at the WAPOR regional conference on 'Elections, News Media and Public Opinion' in Pamplona, Spain, November 24–26. Accessed online at www.zetterberg.org/Lectures/l041115.htm on October 1, 2006.

FURTHER READING

Mail and Internet Surveys: The Tailored Design Method by Dillman (2007 edition) and Survey Errors and Survey Costs by Groves (1989). Each of these texts focuses on the entire process of designing and fielding surveys, not just sampling.

Conducting Research Surveys via E-Mail and the Web by Schonlau, Fricker and Elliott (2002) 'is a practical and accessible guide to applying the pervasiveness of the Internet to the gathering of survey data in a much faster and significantly less expensive manner than traditional means of phone or mail communications.' (Midwest Book Review, 2003)

'Review: Web Surveys: A Review of Issues and Approaches' by Mick P. Couper (2000) published in *The Public Opinion Quarterly*, is an excellent and highly cited article that emphasizes many of the points and

ideas discussed in this chapter. It also provides additional examples to those presented in this chapter.

Sampling Techniques by Cochran (1977) is one of the classic texts on the mathematical details of survey sampling, covering a wide range of sampling methods applicable to all types of survey effort.

Online Panel Research: A Data Quality Perspective by Callegaro et al. (2014) is a detailed examination of and exposition about the current state of online panel research methodology. The text nicely summarizes the what is currently known about the quality of data obtained via online panels, and it discusses the various sorts of errors and biases that can affect online panel results.

Research Synthesis: AAPOR Report on Online Panels by Baker et al. (2010) summarizes the results of a report commissioned by the American Association of Public Opinion Research (AAPOR) Executive Council with the charge of "reviewing the current empirical findings related to opt-in online panels utilized for data collection and developing recommendations for AAPOR members" (Baker et al., 2010: 712). The authors are recognized survey experts from a variety of fields and disciplines and the resulting report is an authoritative treatment of the current state of the art in using online panels for surveying.