Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

The promise and challenge of pushing respondents to the Web in mixed-mode surveys

by Don A. Dillman

Release date: June 22, 2017

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at STATCAN.infostats-infostats.STATCAN@canada.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-877-287-4369

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard table symbols

The following symbols are used in Statistics Canada publications:

. not available for any reference period
.. not available for a specific reference period
... not applicable
0 true zero or a value rounded to zero
0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p} preliminary
^{r} revised
x suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E} use with caution
F too unreliable to be published
* significantly different from reference category ($p < 0.05$)

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, June 2017
Vol. 43, No. 1, pp. 3-30
Statistics Canada, Catalogue No. 12-001-X

The promise and challenge of pushing respondents to the Web in mixed-mode surveys

Don A. Dillman[1]

Abstract

Web-push survey data collection that uses mail contact to request responses over the Internet, while withholding alternative answering modes until later in the implementation process, has developed rapidly over the past decade. This paper describes the reasons this innovative mixing of survey contact and response modes was needed, the primary ones being the declining effectiveness of voice telephone and slower than expected development of email/web only data collection methods. Historical and institutional barriers to mixing survey modes in this manner are also discussed. Essential research on the use of U.S. Postal address lists and the effects of aural and visual communication on survey measurement are then described followed by discussion of experimental efforts to create a viable web-push methodology as an alternative to voice telephone and mail response surveys. Multiple examples of current and anticipated web-push data collection uses are provided. This paper ends with a discussion of both the great promise and significant challenge presented by greater reliance on web-push survey methods.

Key Words: Surveys; Mixed-mode; Web-push; Web; Mail; Telephone; Address-based sampling; Visual communication; Response rates; Measurement differences.

1 Introduction

A surprising, but critical, development in survey design during the early 21st century is the extensive use of web-push data collection methods, i.e., the use of postal mail to obtain questionnaire responses from general public samples mostly over the Internet instead of paper questionnaires. Web-push methods are now being used as a replacement for paper mail-push procedures whereby an attempt is made to obtain responses by mail before using other modes of response such as telephone or personal interviews. Web-push methods are now being used in official government surveys and as a replacement for random digit dialing (RDD) voice telephone surveys.

For example, the American Community Survey, which serves as the major source of state and regional information on U.S. households, began using a web-push approach to data collection in 2013 that includes the possibilities of responding later in the implementation process by mail, telephone or in-person interview. Plans are now in place to use such a methodology for the 2020 U.S. Decennial Census. Web-push data collection with an initial mail request is also being used worldwide. Examples include the 2015 Japanese Census (Statistics Japan 2015), and the 2016 Censuses in Canada (Statistics Canada 2016) and Australia (Australian Bureau of Statistics 2016). Other examples of web-push procedures are household surveys in Switzerland (Roberts, Joye and Staehli 2016) and the United Kingdom's Community Life Survey (United Kingdom Cabinet Office 2016), which is being transitioned from a personal interview mode. In addition, the U.S. College Graduates Survey, conducted every 2-3 years by the National Science Foundation, has completed the shift from mail and telephone data collection to a web-push approach followed by the other two modes of data collection (Finamore and Dillman 2013). These examples are only a few of the major survey efforts around the world that are now using this methodology.

1. Don A. Dillman, Washington State University. E-mail: dillman@wsu.edu.

Use of web-push data collection methods has been encouraged by a number of considerations, ranging from seemingly unfixable problems of RDD telephone surveys to the fact that Postal Service residential address lists or country-wide registration lists now provide the most complete coverage of households. In addition, there are no acceptable ways of drawing probability samples of household email addresses as a means of household contact. Even if email addresses could be sampled, it is not likely that reasonable response rates could be obtained through email-only contact (Lozar, Bosnjak, Berzelak, Haas and Vehovar 2008).

The current heavy reliance on mail contact is surprising, despite the demonstrated potential of mail surveys for obtaining reasonable response rates in the late 20th century (Dillman 2000). Until recently, mailing address sample frames have been mostly unavailable and inadequate. In addition the general availability of a telephone alternative prior to the late 1990's meant that mail was infrequently used for government surveys, with the exception of official government censuses.

My purpose in this paper is first, in Section 2, to discuss the reasons that web-push survey methodologies have been developed and adopted worldwide. Secondly, in Section 3 and Section 4, I describe research efforts that have not only made web-push methodologies feasible, but are improving the effectiveness of such methods in producing reliable estimates of the opinions and behaviors of survey populations throughout the world.

This research has shown, see Section 5, that web-push methodologies are quite promising with regard to improving coverage and response rates, while reducing measurement differences across modes as well as total survey costs. It has also shown that there are many perils that threaten their use, ranging from respondent trust of the Internet to the plethora of devices now available for responding to such surveys, see Section 6. My focus in this paper is to present the substantial promise and many challenges associated with web-push methods for conducting sample surveys. Section 7 presents a summary and a conclusion.

2 Why Web-push data collection is needed

Fundamentally, making contact with households or individuals by one mode, such as mail or telephone, to request that they respond by another mode, is not an ideal data collection procedure. There is bound to be some friction between receiving a postal letter or phone call and then having to go to a different response mode. The switch by itself is likely to take a toll on response rates. Thus, it is not surprising that difficulties in conducting single mode telephone and e-mail/web surveys are the fundamental reason for seeking an alternative.

2.1 The declining effectiveness of telephone surveys

In the mid-20th century, most methodologists considered face-to-face interviews as the only acceptable means for conducting sample surveys (e.g., Parten 1950; Kerlinger 1965). In addition, sampling and surveying households was slow and costly, and therefore limited mostly to conducting national and other large area surveys.

Although telephone surveys had been used occasionally to support data collection (Nathan 2001), development of the telephone as a sole means of collecting survey responses did not occur until the early 1970's, a process described in detail by Nathan (2001). The first three books on methods for conducting telephone surveys appeared in rapid succession, developing marketing (Blankenship 1977), state and special population (Dillman 1978) and national (Groves and Kahn 1979) population survey perspectives. The use of telephone data collection methods advanced rapidly because of the expanding presence of telephone in households and development of the Mitofsky-Waxberg procedure for using random digit dialing methods of selecting households. In addition the declining costs for long distance calling resulted in RDD voice telephone surveys replacing most in-person interviewing (Dillman 2005).

Between 1997 and 2012 the Pew Research Center (2012), a major conductor of social surveys by telephone in the United States, reported declines in RDD response rates from 35 percent to about 9 percent. More recently, Dutwin and Lavrakas (2016) conducted an analysis of telephone response rates for nine organizations. They found that landline response rates declined from 15.7 percent in 2008 to 9.3 percent in 2015, while cell phone response rates declined during this period from 11.6 to 7.0 percent. They also reported that this roughly 40% decline in response is less the result of an increase in refusals than it is an increase in no answers and answering machines of 10 percentage points for landlines and 24 percentage points for cell phones.

However, these results present only the tip of the iceberg with regard to what is happening to telephone. The telephone has changed from being a household device, or landline, shared by all household members to a wireless individually possessed instrument, easily transportable from place-to-place. In the United States, half of all households and 60% of those with children are now wireless only (Blumberg and Luke 2017). At the same time, the presence of cell and/or landline phones in households has reached an all-time high of at least 95% in most European countries (Mohorko, de Leeuw and Hox 2013) and 97% in U.S. households (Blumberg and Luke 2017). One implication of the increased proportion of wireless phones is that household sampling has become much more difficult. It is possible to include mobile numbers in RDD sample frames. However, it has also become necessary to devote precious interview minutes to ascertaining a range of information including number and type of phones in a household in order to determine household selection probabilities.

In addition, one needs to learn whether the person who answers the phone is an adult, and select an appropriate respondent. Also, the landline "inconvenient time problem" of the respondent being interrupted while, for example, fixing dinner and not having time to talk has been expanded to needing to find out if a person who answers the phone is driving a car or engaged in another task where safety emerges as a serious issue. The inclusion of such items takes away from the ability to ask other questions in phone interviews for which considerable pressure exists to keep the length to only a few minutes. In sum, a major effect of the changes in how telephones are owned, regulated, and used has made its use for important data collection efforts, increasingly difficult.

Landline and cell phones jointly face a larger challenge. Fewer and fewer people engage in voice conversations by telephone. This is a huge change from the time when essential communication for business

discussions, maintaining social relationships, and coordinating daily activities in a timely way, were done mostly by voice telephone. Email and texting have largely replaced that use. Talking over the phone with a survey interviewer is increasingly out of sync with other aspects of people's daily lives.

Answering machines now take most incoming voice calls on both landline and cell phones. Not answering one's phone is no longer considered rude. Desired calls from children and other close relatives may be assigned a special ringtone to draw the call recipient's attention. Phone calls from specific numbers can also be blocked, or, on smartphones, swiped away. In addition, both landline and cell telephone numbers are now transportable across type of phone and area codes in the United States and different federal rules apply to automatic dialing of phones.

Yet, another emerging problem with telephones is that the repeated contacts necessary for achieving reasonable response rates for all types of phones are becoming less effective. Increasingly, telephone interviewers have only one chance, lasting just a few seconds, to persuade people to be interviewed. The appearance of the originating telephone number and/or source on caller identification screens makes it increasingly likely that follow-up phone calls can be avoided. Also, the plethora of marketing and fundraising calls has produced an environment in which fewer people are willing to answer the telephone, let-alone be interviewed. An additional challenge associated with cell phones is that their use has brought with them a greater likelihood that requests to be surveyed come when the call recipient is in the midst of business and work activities that are not conducive to people taking time to be interviewed.

The decrease in RDD telephone interviews was slowed down for awhile by research that has shown intensive callbacks to increase response rates does not improve the accuracy of results (Keeter, Miller, Kohut, Groves and Presser 2000) and other research that has suggested that occurrence of non-response error (the differences between respondents and non-respondents) is not closely related to response rates (Groves and Peytcheta 2008). The telephone's continued use was also encouraged by the great investment that organizations have in telephone hardware, software, and specialized personnel, many of whom had not done other types of survey data collection. However, the continued decline of telephone response rates in recent years noted by Dutwin and Lavrakas (2016) and measurement concerns have reduced the credibility of doing stand-alone telephone surveys.

2.2 The slower than anticipated emergence of email/web only surveys

In the mid 1990's when telephone response rates were starting their persistent decline, Internet surveying, the expected replacement, was beginning its rapid development (Dillman 2000, Chapter 11). Yet, two decades later, its use for general population surveys remains limited.

Household Internet penetration in the U.S. and many other developed nations now exceeds 85%, which is higher than for telephone when the rapid development of surveys by telephone occurred in the early 1970's (Nathan 2001). The lack of Internet in some households (e.g., 41% of U.S. adults 65+ and 26% of individuals with only a high school or less education) remains a concern (Anderson and Perrin 2016), but each year sees that becoming less of a problem. Internet use skills are now fundamental to the educational process, to organizational operations, and to accessing consumer services. Yet, the barriers to obtaining web responses for household surveys when using only email contacts remains huge.

There is no household or general population sampling algorithm for email addresses that will provide a known non-zero chance of being selected for survey participation, as calling random numbers has provided for telephone surveys. Email addresses do not exist in standard formats as is the case for our 10 digit telephone numbers that identify an area code, exchange, and the 10,000 number possibilities in each exchange. People within households are also likely to have multiple email addresses so that the probabilities of reaching specific households or other sample units cannot be calculated. In addition, some of the population that are most computer literate – young adults – have developed a reputation for minimizing their use of traditional email systems. Instead they focus heavily on Facebook, Snapchat, and other instant messaging applications for connecting with friends and acquaintances.

In addition, web response rates for random samples of existing email addresses are likely to be as low or lower than those achieved for today's telephone surveys (Lozar et al. 2008). And, they are likely to include disproportionately high numbers of individuals who are younger and better educated, despite the fact that many younger people rely on other ways of connecting electronically that make them only occasional users of traditional email. People's computer inboxes are typically a crowded space, with unsolicited and unwanted emails being more prevalent than the number of unwanted telephone calls once was. In addition, emails are often scanned and deleted based solely upon source or after reading only a few beginning words of the accompanying message.

Changing computer technologies are also contributing to web survey nonresponse. Smartphones that fit in one's purse, handbag, or pocket, now have far more computer power than desktop computers had when internet surveying began (e.g., Friedman 2016). Their constant presence with people has led to these devices being used as the first responders for scanning and discarding unwanted requests. Some users may defer answering survey requests until they get to a laptop or desktop with a full-scale keyboard. However, for some individuals, smartphones are now the dominant, or even only, device for responding to all emails.

When our dominant survey mode was telephone, interviewers could usually focus the respondent's attention on survey questions and guide that person through the interview. On desktops, laptops, and now tablets that are used in a person's office or home, considerable mental concentration by the respondent can often be achieved. In the smartphone era when people are as likely to be on the move from one place to another, concentration on survey content seems somewhat less likely to be achieved. It is evident that the proportion of surveys completed over smartphones is increasing as a proportion of all web completions (Couper, Antoun and Mavletova 2017). However, there appears to be no evidence that the smartphone delivery of web survey requests is increasing total survey response, and may in fact be lowering it. In addition, breakoff rates are much higher for smartphones than desktops and laptops.

Concern about the consequences of attempting to answer an electronic survey is another factor limiting the potential effectiveness of email/Internet surveying. The ease and low cost of sending out massive numbers of email survey requests, has increased the likelihood that people receive requests from organizations that they know nothing about. In addition, considerable fear exists that such requests may be originating from sources that are imitating legitimate sponsors, and attempting to deliver malware, ransomware, and/or collecting data that can be used for other nefarious purposes. Thus, people who are willing and able to respond to legitimate web surveys may be unwilling to take that risk. For many, the internet is a scary place where a "consumer beware" climate prevails.

For all of these reasons, it is hardly surprising that low cost email contact/web response surveys have not become the method of choice for conducting random surveys of the general public needed for public policy purposes. Even if the challenge of drawing probability samples could be solved, multiple issues including computer technologies, the circumstances in which potential respondents encounter survey requests, and mistrust about who is requesting a survey request and how data might be used, are limiting its ability to replace the telephone.

3 Overcoming barriers to the acceptance of mixed-mode designs

3.1 Historical barriers to mixing modes

Use of more than one survey mode, as a means of contacting and/or asking questions, was seldom done in the late 20th century. Gaining acceptance of mixed-mode designs for any purpose has been a slow process. The biggest barrier prior to the 1990's was simply the lack of perceived benefit. Response rates to in-person, telephone and mail surveys were considered high enough that use of a second or third mode was considered unnecessary. A significant exception was those surveys in which less expensive survey methods were used early in the data collection process, but in-person methods were necessary to achieve response rates over 90%. The U.S. Decennial Census from 1970-1990, which followed a mail questionnaire start with in-person and in some cases telephone follow-up is an example.

Another barrier to the mixing of survey modes was that the data collection technology of the times made it difficult to simultaneously implement multiple modes of data collection in a single survey. The lack of networked computers and software meant that using a second mode of data collection required finishing up data collection for one mode before switching the effort to a separate data collection unit for implementation of a second mode (Dillman, Smyth and Christian 2009, Chapter 8). An earlier review of the use of telephone in mixed-mode surveys in the late 1980's found that few mixed-mode surveys had been done, other than pre-letters to an anticipated telephone or in-person interview (Dillman and Tarnai 1988).

During the 1990's it became apparent that new methods of surveying needed to be developed. Response rates, especially for personal interviewing and telephone, had begun to decline (Brick and Williams 2013). Coverage problems were also increasing, as locked multi-unit buildings and gated-residential communities made it impossible to reach many households in-person. The landline coverage of households also began its long inexorable decline that now leaves about half of U.S. households without such connections.

Interest in coordinating the use of multiple modes in some way to improve response rates brought attention to interview concerns that had previously been ignored because of the practical barriers to mixing modes. For example, interviewed respondents often gave socially desirable responses so that estimates of desirable behaviors of e.g., "having voted in the last election" were higher than the actual behavior. In addition, estimates of undesirable behaviors, e.g., smoking marijuana or having sex outside of marriage were lower (de Leeuw 1992). Differences were also being observed between mail answers to survey questions on the one hand vs. telephone and in-person surveys where respondents gave more extreme positive answers on opinion questions (de Leeuw 1992; Tarnai and Dillman 1992). Research had also

suggested that respondents were more likely to choose the first response categories in mail surveys, a primacy effect, and the last categories in telephone surveys, described as a recency effect (Krosnick and Alwin 1987). As a consequence, it became increasingly difficult for survey sponsors to argue that telephone, or even in-person interviews, were superior survey modes.

Mixed-mode surveys were proposed as a potential, albeit imperfect, solution to the problems of individual survey modes. Five types of mixed-mode surveys were identified, ranging from collecting the same data from different members of a sample to using one mode only to prompt completion by another mode (Dillman 2000, page 219). The major advantage of combining modes appeared to be improvements that could be achieved in coverage and response rates. The major difficulty identified was the prospect of measurement differences when different response modes were used.

A critical article by de Leeuw (2005) influenced an important transition in thinking about mixed-mode surveys. She articulated a variety of accepted possibilities for combining survey modes and provided evidence that an increase in use of mixed mode surveys was occurring. She also noted the transition that occurred from debate on which survey mode was best for a particular study, to how modes could be used together and produce better results.

A contextual change was also underway as modern societies throughout the world began shifting from person-mediated activities (e.g., getting money from tellers in banks, making travel reservations through an agent, and purchasing goods in stores and from catalogues) to self-administration (Dillman 2000). But researchers had not yet answered the question of whether interviews by telephone could persist in the face of these self-administration trends.

3.2 Institutional barriers to the joint use of survey modes

Considerable reluctance, if not outright opposition, existed with regard to mixing survey modes and especially to giving up cherished ways of asking questions differently in different survey modes (Dillman 2000). One of the consequences of the emergence of new ways of collecting survey data in the last third of the 20th century is that data collection staff became quite specialized. Some organizations conducted surveys by only one mode. It was common for some data collection staff and their organizations to do only telephone surveys, and to a lesser extent mail surveys. A few large firms had in-person sampling and data collection units. A tendency existed to want to do surveys by the mode that a group knew best. This tendency was exacerbated in the late 90's as Internet-only survey organizations began to emerge.

In addition different styles of wording questions by individual modes had emerged. Interviews tended to withhold "don't know" categories, offering them only when respondents would object. Designers of paper and web surveys often used "mark all that apply" question formats to make responding easier, but the awkwardness of that format on telephone led to using only forced choice formats that obtained an answer after each individual item was presented on the telephone. The issue facing survey designers was whether to maximize question formats for the mode, or try to keep the same stimulus across all modes (Dillman and Christian 2005).

One of the factors supporting this tendency to stick with what surveyors knew best was the recognition that single mode surveys were best for many survey situations. That situation existed for each of the modes.

In-person interviews were the only way of obtaining adequate coverage for certain national surveys such as the Current Population Survey that produces employment rate estimates. RDD telephone surveys were the best means of conducting election and other cross-sectional household surveys. Mail was the most adequate means when one was conducting regional and local surveys for which only residential addresses were available. Interactive Voice Response surveys were most practical for many customer satisfaction surveys when people contacted calling centers to obtain a particular service. And Internet surveys quickly became the go-to methodology for client surveys and other situations for which email addresses had been previously collected.

This situation marked the development of interest in "tailored design," i.e., recognition that different modes of data collection fit better with surveys of particular populations, survey topics, and data collection situations. This trend in survey design now persists more powerfully than it did at the turn of the century. It is clear that picking the single best survey mode for a particular survey is increasingly inadequate because of negative effects on coverage, response rates and nonresponse error.

As the 20th century came to an end, there was much uncertainty with regard to where data collection methods might be headed. Prospects seemed dim for continuing sole reliance on either in-person or telephone surveys. The coverage challenges and costs were growing significantly, and it seemed unlikely that response rates were likely to improve for voice telephone surveys. Great interest existed in replacing interview methods with the Internet, but at the turn of the century only about half of the households in the U.S. as yet had computers, and even fewer had access to the internet (Dillman 2000).

4 Development and testing of web-push mixed-mode data collection

In the first decade of the 21st century, the idea of using multiple survey modes to contact individuals and obtain survey responses seemed to be an issue whose time for serious in-depth exploration had arrived (e.g., Tourangeau 2017; de Leeuw, Villar, Suzer-Gurtekin and Hox 2017). In addition, the same information technologies that brought on the internet carried with it the potential for managing the simultaneous use of multiple modes of data collection effectively as well as efficiently, thus removing the primary practical barrier to conducting mixed-mode surveys.

The effective development of web-push methods meant addressing multiple issues all at once in order to learn whether such an approach would be effective. These issues ranged from responding to household coverage problems and developing an understanding of how visual communication differed from aural communication, as well as expanding our theoretical thinking about what influences people to respond to survey requests.

A major, unanswered question was whether self-administration could replace human interviewing, and would the results be better or worse. A confounding challenge was that survey modes varied significantly in individual coverage, response rates, and bias in how people responded to the use of them, with each survey mode perhaps being better in some situations and worse in others.

4.1 U.S. Postal Service residential address lists now provide excellent household coverage

Because the U.S. Postal Service makes available, through vendors, complete residential address lists, it is possible to send mail requests to nearly all residences in the United States (Harter, Battaglia, Buskirk, Dillman, English, Mansour, Frankel, Kennel, McMichael, McPhee, Montaquila, Yancey and Zukerberg 2016). These computerized residential lists are provided without names, just as RDD telephone lists do not have names. The lack of names is not a barrier to obtaining response from households, as shown by a series of response rate studies using U.S. Decennial Census address lists (Dillman 2000, Chapter 9). In addition it does not direct a mailing to only one person in households whose occupants are now less connected to one another than when marriage rates were higher. It may also allow more accurate respondent selection by not having to overcome the limitations of mailings being associated with just one household member.

One of the first large-scale studies to evaluate the use of address-based sample (ABS) with postal data collection was by Link, Battaglia, Frankel, Osborn and Mokdad (2008). It found for a 2005 Behavior Risk Factor Surveillance System (BRFSS) questionnaire that a mail questionnaire sent to an ABS sample obtained significantly higher response rates than those obtained by RDD in five of the six states surveyed. The authors concluded with appropriate caution that the true potential of ABS might be in facilitating mixed-mode surveys that also involved telephone follow-up, and provided a strong recommendation for further study.

Other research at this time showed that ABS samples had very high coverage which was improving as city-style addresses were replacing less specific addresses, such as rural routes (O'Muircheartaigh, English and Eckman 2007; Battaglia, Link, Frankel, Osborn and Mokdad 2008). In addition, a series of studies showed that a two-step ABS mail survey (screening of households for the presence of school children, followed by a detailed questionnaire on a particular child) produced better results than a two-step RDD approach, with significantly higher response rates (Brick, Williams and Montaquila 2011; Williams, Brick, Montaquila and Han 2014).

These studies contributed significantly to establishing the high coverage qualities of address-based sampling as an alternative to RDD sampling. However, they stopped short of testing the possibility that the contacted households could be persuaded to respond by web to mailed requests.

4.2 Identifying and overcoming measurement differences between visual and aural surveys

A quite different concern that limited interest in address-based sampling with paper and/or Internet questionnaires was that people's answers to questions were likely to be different from telephone responses. There were two aspects to this concern. The first was that without an interviewer, respondents could not be given extra encouragement when they were unable or reluctant to answer a question, nor could misunderstandings of questions be corrected. The second was the long-standing evidence that social desirability and the tendency to agree (acquiescence) were greater for telephone than self-administered (mail) survey responses (de Leeuw 1992). Traditionally, the benefit of having an interviewer present was viewed as outweighing the potential bias from the latter.

A survey sponsored by the Gallup Organization in 1999 provided a new perspective on these differences. This test revealed that asking people to respond in an interview to aurally received stimuli (either by voice telephone or Interactive Voice Response) produced similarly more positive answers than those given to visually delivered stimuli, either by mail or Internet questionnaire (Dillman 2002; Dillman, Phelps, Tortora, Swift, Kohrell, Berck and Messer 2009).

Discoveries in how visual information is processed reported by Palmer (1999), Hoffman (2004) and Ware (2004), provided theoretical insights into the separate actions taking place as the eye takes in the information and the brain processes it to make sense of what is on the page or screen. Application of these concepts provided an understanding of the reasons that self-administered questionnaires often produced different answers than interview surveys, as revealed by the Gallup study. Respondents are guided through visual questionnaires by multiple languages that communicate meaning. They include symbols, numbers and their graphical composition (size, spacing, color, symmetry, regularity, etc.) that affect how information on paper and web pages is navigated, mentally grouped and interpreted (Dillman 2007, pages 462-497; Tourangeau, Couper and Conrad 2004). Additional research showed that compliance with branching instructions could be improved dramatically through changes in symbols, font size, font brightness (Redline and Dillman 2002; Christian and Dillman 2004), and the placement of those branching instructions in relation to answer choices (Redline, Dillman, Dajani and Scaggs 2003; Dillman, Gertseva and Mahon-Haft 2005).

Another major cause of measurement differences across modes became apparent: questions were often worded differently for each mode and presented using different structures (Dillman and Christian 2005). For example, researchers had a long tradition of asking forced choice questions individually on telephone surveys when surveying people's opinions on a list of items, but they often converted it to a check-all reply format for items presented as a group on mail questionnaires (Smyth, Dillman, Christian and Stern 2006). This practice was carried over to web surveys. New research showed that using forced-choice formats on both visual and aural modes would bring respondent answers much closer together (Smyth, Dillman, Christian and McBride 2009). Research also showed that open-ended questions to mail and web surveys would be comparable if similar visual construction was used for both mail and web (Smyth, Christian and Dillman 2008). In addition, it was learned that variations in scalar question formats (e.g., fully labeled vs. polar point labeled) produced dramatic differences in answers within visual modes (Christian, Parsons and Dillman 2009).

Unified mode construction – the use of the same wording and visual layout of survey questions – was proposed as a way of removing measurement differences across these modes (Dillman 2000). Unified construction could easily be accomplished for many types of questions (e.g., to present "don't know" categories to all respondents instead of only those who would not choose an offered answer choice), as typically done by telephone interviewers. However, in other instances construction that differed across modes was both practical and would reduce errors, e.g., automatic branching to the next appropriate question on web and telephone. This form of presentation cannot be accomplished with branching items for paper questionnaires where all options have to be printed because of not being able to anticipate how people will answer those items.

The major contribution of unified mode construction has been to reduce concern about measurement differences being encouraged by multiple modes of survey response. An exception is that strong evidence exists that telephone response to opinion scales using vague quantifiers are consistently more likely to produce extreme responses on the positive end of the scale and less use of intermediate categories, than are web and mail questionnaires (Christian, Dillman and Smyth 2008). The apparent reason for this difference is that the visual presentation of intermediate response categories is more visible, and therefore accessible to respondents than when those same categories are read over the phone, a process that makes the end categories more prominent in respondent minds, Dillman and Edwards (2016).

Another difference that unified mode construction does not resolve is how people answer socially desirable questions. However, self-administered (visual) questionnaires are generally thought to produce more honest answers.

The accumulation of research on visual vs. aural design issues has provided survey designers with crucial tools, the use of which partly eliminates measurement differences that might undermine coverage and response benefits of mixed-mode surveys. The practice of unified mode design was crucial for initial development and testing of the web-push methodology described below.

4.3 The sequential development of an effective web-push methodology

A sequence of ten tests of web-push data collection procedures was conducted by a team of researchers at Washington State University between 2007 and 2012 in five separate data collections. The plan that guided these experiments was to build upon what was learned from the initial tests to design and implement the later tests. All experimental comparisons used the equivalent of 12 page paper questionnaires, containing 50-70 numbered questions, requesting 90-140 potential answers. They were designed to be the equivalent of 20-30 minute interview questionnaires. The studies were on a variety of topics – community involvement and satisfaction, use of information technologies, economic and social effects of the 2008 recession, energy use attitudes, and understanding water quality and management. Researchers varied the topics in order to reduce concerns about the effect of topic on response rates and data quality.

The populations surveyed ranged from a rural region of Idaho and Washington and statewide surveys of Washington, Pennsylvania, and Alabama conducted from Washington State University, to surveys of Nebraska and Washington residents sent from the University of Nebraska and the same surveys sent to both states from Washington State University. Implementation procedures varied, but included from 4-5 mail contacts, with the mail-back questionnaire option provided in either the 3rd or 4th contact. A small token cash incentive was sent with the initial response request, and in some instances a smaller incentive was sent with the paper questionnaire when it was withheld until the 3rd or 4th contact. Detailed procedures for each of the studies are provided elsewhere (Smyth, Dillman, Christian and O'Neill 2010; Messer and Dillman 2011; Messer 2012; Edwards, Dillman and Smyth 2014; Dillman, Smyth and Christian 2014).

The initial test in a rural region of Idaho and Washington resulted in 55% of households responding to the web-push treatment, with 74% of those responses coming over the Internet. This test also revealed that enclosing a paper questionnaire and offering an immediate choice of modes produced a significantly higher

response rate of 63% (Smyth et al. 2010). Unfortunately, nearly 80% of those responses came by paper, too many to warrant the cost of setting up the web data collection. Because of that effect and the initial promise being shown of getting more than half the households to respond over the Internet in the web-push treatment, experimentation on the choice methodology was discontinued. We also found from this initial test that a paper-push treatment that withheld offer of a web option until the last contact produced only two percent of the responses over the Internet. Based upon this result the web follow-up was discontinued after two additional tests with similar results. In addition, results from this initial rural region study encouraged us to carry forward the web-push with paper follow-up for additional testing with statewide populations.

Across all ten experiments that involved five states, the web-push surveys produced a mean response rate of 43%, ranging from 31-55% (Figure 4.1). The mail-only comparisons produced a mean response rate of 53%, with a range of 38-71%. On average, 60% of the responses to the web-push treatments came over the Internet. Experimental treatments in one of the studies showed that the incentive enclosed with the web-push request improved the web response dramatically, from 13% to 31%, or about 18 percentage points (Messer and Dillman 2011). Although an RDD comparison was not included in any of the experiments, the results from the web-push procedures were undoubtedly much higher than would have been obtained by telephone for these long questionnaires, had such a comparison been included.

An item nonresponse comparison was made for three of the experiments to determine whether the mail follow-up questionnaires obtained higher item nonresponse rates than the web responses obtained in those treatment groups. For the regional study in 2007 and two statewide studies in 2009 the follow-up paper questionnaires produced item non response rates more than twice as high, 8.2% vs. 3.6% for those who responded by web. However, when the overall item nonresponses for the web-to-push treatment groups (web plus mail responses) were compared to mail only treatment group responses, there were virtually no differences between groups, being 5.3 and 5.7 respectively. The authors speculated that the initial web responses were being provided by “better” respondents, while the later responses by mail were generating responses from less able respondents, as indicated by being older and having less education. (Messer, Edwards and Dillman 2012).

Populations likely to be unfamiliar with Washington State University – the sponsor of these studies – had significantly lower response rates, especially among those responding on the web. For example, only 12% and 11% responded in Pennsylvania and Alabama respectively, compared to 28% in the Washington survey (Messer 2012). A water management study conducted by the University of Nebraska and Washington State University provided additional insight on this phenomenon by sending requests for responses to households in the other state. The web-push response was 6.1 percentage points lower among Washington residents and 14.7 percentage points lower among Nebraska residents when surveyed from the University located outside the state (Edwards et al. 2014). Virtually all of this decline occurred in the internet responses which decreased from 32 to 26 percent in Washington and from 38 to 23 percent in Nebraska, when the response requests came from the opposite state’s university. We speculated that responding over the Internet is more sensitive than mail responses to the lack of familiarity and trust of the survey sponsor.

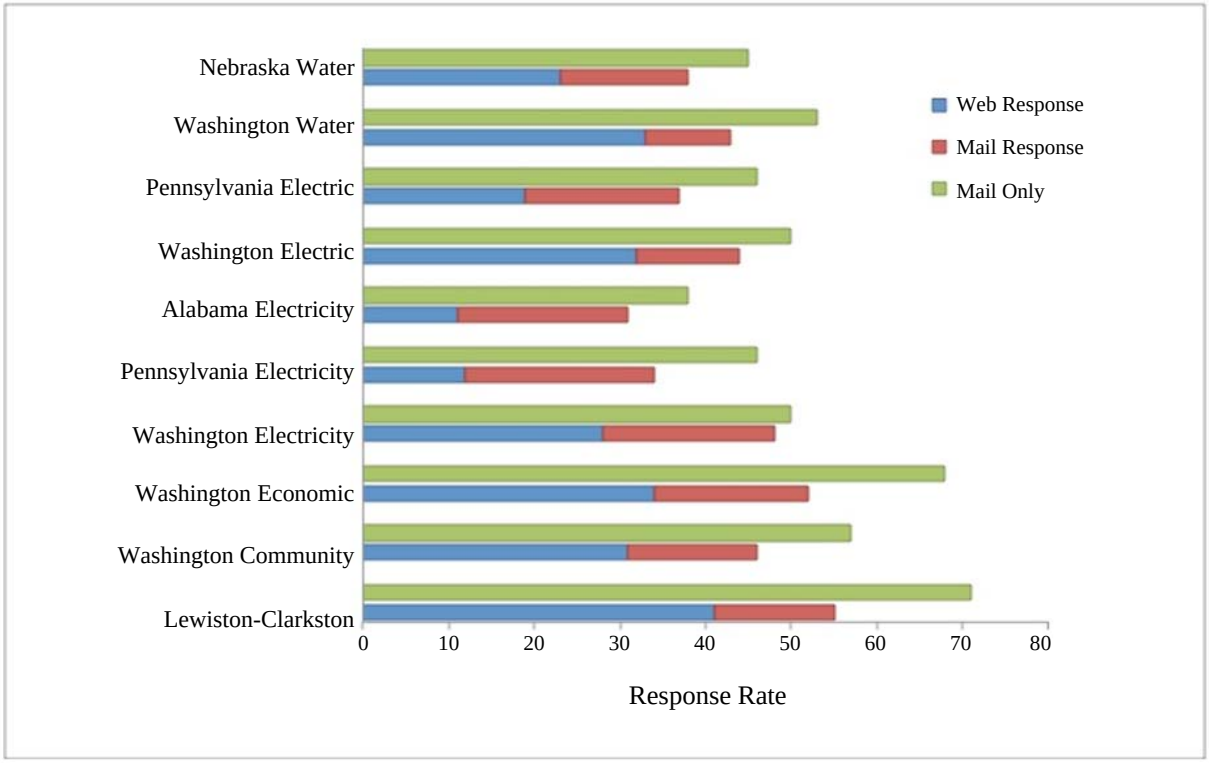

Figure 4.1 Mail-only treatment response rates vs. web-push response rates with proportion of that response received by each mode. (Dillman et al. 2014, Chapter 11).

The ten comparisons from these web-push studies revealed that those who responded over the Internet in the web-push treatment groups were significantly different than those who responded later to the mail questionnaire. For example, web respondents were younger, more educated, had higher incomes, and were less likely to live alone (Messer and Dillman 2011). However, the combined web and mail respondents in the web-push treatments were demographically quite similar to the mail-only treatment groups. The research concluded that individuals prone to respond by web could also be convinced to respond to the mail-only treatment. This finding was reinforced by the fact that a paper questionnaire follow-up to the web-only request produced significant improvement in response rates, whereas a web follow-up to a mail-only approach did not produce Internet responses that significantly improved overall response.

Although the web-push and mail-only treatment group responses were quite similar, the unweighted data exhibited nonresponse bias for certain demographics. Respondents had more education and children in the home than those who responded to the American Community Survey (discussed in more detail below) that now uses mail, web, phone and in-person interviews to obtain close to 97% response rates, and is relied on for producing official U.S. statistics for all U.S. states. Such comparisons were beyond the purpose and scope of these experiments and more investigation needs to be done to understand the nature of such differences. In addition, costs per respondent were not shown to be lower for Internet responses, because of contact costs being about the same for web-push and mail only methods, while producing fewer respondents (Messer and Dillman 2011). That seems likely to change as use of the Internet continues to expand to more people and areas of life.

Overall, the outcome of this coordinated set of studies made it clear that the web-push methodology offers considerable promise for obtaining web responses to household surveys. It was also clear that paper follow-up questionnaires would improve representation of people unable or unwilling to respond over the Internet.

4.4 Additional web-push tests on other populations and situations

In recent years the use of web-push data collection strategies has expanded and they are widely used in government, university, and private sector surveys in multiple countries. Uses have also spread beyond general public populations, and now involve survey situations where requests to respond via the Internet are not limited to mail contact. In addition, some surveys involve as many as three modes of contact and three modes of response, with the intent of getting very high response rates, while pushing as many respondents as possible to the web, in an effort to lower survey costs.

In 2013, the American Community Survey was converted from a sequence of mail-telephone-in person requests, to beginning with an Internet response, followed by the three remaining contact and response procedures (United States Census Bureau 2014, Chapter 7). The law requires U.S. citizens to respond to the American Community Survey (formerly the long form in the Decennial Census). Therefore, the overall response rate for occupied households was about 97%. Tests of web-push strategies began in 2011, when an initial experiment confirmed that web-push resulted in dramatically higher response rates (28% vs. 10%) over the Internet than did a "choice" strategy that also offered mail in the first contact (Tancreto 2012). In 2013, 28% of the responses from occupied households occurred over Internet, 22% by mail, 6% by telephone, and 43% by in-person interview. Thus, about 51% of the self-administered responses were over the Internet, a proportion that increased to 58% in 2015. Tests are now underway in support of plans to convert the 2020 Decennial Census to web-push methods with a similar follow-up.

The Japanese Census was converted to a web-push methodology in 2015 (City of Sapporo 2015). On-line response was about 37%, with one third of those responses coming from smartphones, which are used extensively in Japan. The remainder of the response was obtained by mail questionnaires and enumerator visits. The 2016 Australian Census and 2016 Canada Census were also conducted using web-push methodologies. Although final results are not yet available, it is known for Canada that 68% of households responded over the Internet, 20% by mail, with an additional 10% through enumerator visits, for an overall response rate of 98% (Statistics Canada 2016). The proportion of Internet responses in the Canadian Census is the highest I am aware of for a web-push household survey. In some areas of Canada a paper questionnaire was included with the request, providing a choice of response modes to respondents. The high Internet response (68%) and Internet plus mail response (88%) suggests great promise for use of a web-push methodology in that country and perhaps others with high Internet penetration.

A newly developed National Child Health Survey – developed as a replacement for a previous RDD household survey in the United States – plans to screen an address-based sample of children, then select a child for detailed reporting of health issues. However, instead of using two separate mail data collections, they tested in 2015 the possibility of reducing the process to do a one-step, in which the computer uses study criteria to immediately select and administer a topical health questionnaire for one child. This procedure

seeks to improve upon the two-step mail-only response process recently developed for the National Child Education Survey. Results from a pretest in 2015 were promising and it is now going through a second stage of testing.

The U.S. Residential Energy Consumption Survey, conducted for many years by the Energy Information Administration through in-person household interviews, is in the process of being changed to a web-push survey. This survey is noteworthy because it combines a cash incentive with the initial request to respond over the web, and also provides a post-incentive. The post incentive was deemed especially important because of the cost savings it provided by not having to send in-person interviewers to nonresponding households (Biemer, Murphy, Zimmer, Berry, Deng and Lewis 2015).

Not all web-push surveys use address-based sampling. The 2010 National Survey of College Graduates (NSCG) began sampling individuals who reported being a college-graduate in the previous year's American Community Survey and asked them to complete the NSCG, which is conducted every two years (Finamore and Dillman 2013). Postal addresses, as well as telephone numbers, were mostly available for the households where they had lived the previous year. Prior to 2010, households had been selected from the Decennial Census Long Form (last completed in 2000) with telephone, mail, and in some cases in-person interviews. In 2010, comparisons were made among pushing people to the telephone, pushing respondents to mail, and pushing respondents to the web, followed by use of the other two modes. All three of these treatments were followed by a final telephone effort in which responses could be made by that mode or either of the others. Two results were particularly important. First, all three response rates were within a few percentage points of each other, ranging from 74-77%, for this voluntary survey. However, the web-push strategy, in which 53% reported by web, proved to much less expensive, $48 per respondent vs. $66 for mail first and $75 for telephone first. It was concluded that the results from each procedure represented the original sample quite well.

A recent voluntary survey of spouses of U.S. military members compared a web-push strategy with a mail-push strategy. The web-push methodology produced a significantly higher response rate, 33% vs. 28%, with 87% of the web-push responses being received over the Internet (McMaster, LeardMann, Speigle and Dillman 2016). The web push strategy was also much less expensive, $61 per respondent vs. $89.

The success of the web-push strategies for the college graduate and military member studies may be for different reasons. All of the NSCG participants had at least a four-year college degree. Participants in the Family Study of Military Members were also relatively young. The authors of the latter study suggested that the fact that military members rely greatly on the Internet for communication with spouses during deployment might account for its greater effectiveness than mail-push methods.

Many other tests of a web-push methodology have emerged during the past decade. A Swiss study has shown that response rates of about 72 percent of households drawn from Swiss registration lists with 44% by web, 20% by mail and the remainder by telephone or in-person interviews (Roberts et al. 2016). In the United Kingdom, in-person interviews have been relied on far more extensively than telephone for conducting national statistical surveys. Recently, a decision was made to convert the Community Life Survey from an in-person interview to a web-push followed by mail strategy (United Kingdom Cabinet

Office 2016). This decision was made in order to lower costs, while also increasing the sample size. It remains to be seen what results will be obtained.

Private sector uses of web-push methods for specialized survey populations have also evolved. Nexant now conducts surveys of gas and electric utility customers with web-push methods. In the past telephone surveys were the dominant method. Companies whose customers are to be surveyed are able to provide postal addresses and telephone numbers for nearly all customers and email addresses from 20-40% of households (Sullivan, Leong, Churchwell and Dillman 2015). Following a procedure developed by Millar and Dillman (2011), emails are sent to those households to arrive shortly after the letter of request that contains as $2 incentive, followed by another email three days later, and if there is no response another paper survey is sent. Multiple tests have produced response rates of 40-80% with 80-90% of responses who received this email augmentation of the mail contact responding online, compared to about 35-70% of those without email addresses. Responses can be nudged upwards by 8-10 percentage points with a follow-up phone call to those without email addresses, compared to 1-2% for those with email addresses.

5 A promising future, but difficult challenges remain

5.1 Reasons for optimism

The development and deployment of web-push methodologies for survey data collection during the past decade provide reasons for optimism that higher quality survey data collection can be accomplished. That optimism arises less from the excitement over a specific approach of contacting people and convincing them to respond on the web than it does from a combination of considerations.

Address-based sampling now provides excellent household coverage and is conducive to the use of respondent selection procedures. Substantial proportions of survey populations can be contacted by one mode (mail) and encouraged to respond by another (web, or telephone). Survey sponsors not known to the recipient of the request can be legitimized through mail contact in ways that cannot be accomplished with email requests that go mostly unread or voice telephone requests that go mostly unanswered.

Postal contact also allows the sending of small token incentives with the request, thus providing motivation for making the transition from letter to computer and entering a URL (Uniform Resource Locator) and password. Multiple mail contacts provide the opportunity to offer more complete explanations of why a survey is being conducted and how the results will be used. Sending a paper questionnaire alternative in a later contact not only increases responses rates significantly, but brings in types of households not represented well among the initial internet responses. Several studies have also shown that the ability of web-push designs to bring in from half to three-fourths of all respondents quickly over the Internet, depending upon the sample frame and modes of contact, can reduce survey costs.

When email addresses are available for sample units, as is now the case for some survey populations, email augmentation (i.e., the sending of a quick email follow-up to the initial postal request to provide an electronic link that makes it easier for the recipient to respond over the internet) has been shown to improve web response considerably. Similarly, when telephone numbers are available, a telephone augmentation can be effective for improving response. The concept of using such contacts to augment previous mail contacts

encourages surveyors to think not just about stand-alone contacts, but how each contact becomes part of an overall response strategy.

As shown by the American Community Survey, The Canadian Census, National Science Foundation, and Nexant studies, multiple modes of contact and response provides the potential for achieving response rates that many survey sponsors thought were no longer possible. The ability to approach people with repeated requests to respond – and to do so in different modes – improves survey response more than any single mode of contact and/or response.

In addition, relying to a great extent on self-administration (internet and mail) achieves a better cultural fit with people than does a voice telephone conversation, which is increasingly out of sync with routine communication behavior that places great emphasis on texting and email. Also, changes in questionnaire construction methods from using different question structures and wording in each mode in the spirit of creating what's best for each mode through unified mode constructions assists in avoiding measurement differences across survey modes.

Over time, it seems likely that an increasing proportion of adults will be willing and able to respond to surveys over the Internet. Thus, web-push data collection procedures seem consistent with other societal trends that favor the internet over other forms of communications.

The promise of web-push survey methods stems from its ability to reduce survey error from coverage and survey nonresponse. In addition, our greater understanding of how visual vs. aural construction of questionnaires affects answers and the use of unified mode construction methods makes it possible to reduce measurement differences and error. It seems likely that the number of surveys using web-push methods is just beginning.

6 Challenges facing web-push data collection

Despite the potential of web-push data collection methods, there are also uncertainties regarding whether the use of Internet surveying will continue to expand in use. These concerns are the focus of the final section of this paper.

6.1 Fear of responding over the internet

When the push-to-web Australian Census began in 2016, a series of denial of service (DOS) attacks on the site prompted the Bureau of Statistics to turn off the system for fear of hackers. Such attacks are designed to overload a server with traffic, thus making it inaccessible to the intended users. This is only one kind of attack that might be made on a particular survey or computer user. Others include sending malware (e.g., spyware or ransomware) designed to gain access to or damage a computer that users unknowingly access by opening attachments or clicking on links. In addition, phishing emails are sometimes sent. They are designed to trick people into opening them, and providing personal information, for example, by appearing to be sent from a user well known to the recipient. The result of these various possibilities is to cause many people to worry about the security, or lack thereof, of the website, and information they provide in response

to web survey requests. The lack of trust in web surveys and concerns that information could be kept and used for non-survey purposes are also potential barriers to response.

Large scale surveys, especially those that have a great deal of public visibility, such as a nationwide Census that involves widespread prior communication inviting a response, present an inviting target for those hoping to harm the response process. Thus, even though sponsorship is known the perception of risk may be substantial. In the case of the Australian Census, the marketing focus on inviting everyone to respond on a particular "census day," made the situation worse than might otherwise have been the case. Thus, in addition to having to combat the potential of a cyber attack, survey sponsors face the challenge of restoring confidence in the data collection system.

Intentional attacks on individual computers and devices or on specific surveys are probably the largest peril facing surveying that involves the Internet. They are also a major justification for developing multiple response mode opportunities, and not relying entirely on the Internet. The reliance of web-push methodologies on multiple modes of responding provides some measure of protection for attacks on a particular survey, just as it now provides an alternative for those who now consider an internet response unacceptable. In especially large surveys, such as countrywide censuses, shifting away from asking everyone to respond on the same day may also lessen exposure as well as the impact of some of the potential Internet problems.

It is difficult to anticipate whether technological and social control advancements will negate risks associated with computer use. For now, this is an issue that threatens successfully surveying over the Internet that cannot be ignored.

6.2 Smartphones and the purse/pocket problem

A second, but quite distinct issue now challenging Internet data collection is the use of multiple devices for responding. Increasingly, people carry a computer device – mainly smartphones – with them. In most respects this is a very positive development. Because people carry the capability for providing a survey response with them throughout the day, survey requests can be responded to almost anytime from anywhere. This constant availability also brings to the fore what can be described as the pocket/purse problem. There are size preferences and probably limitations on the devices most people are willing to carry with them for use in cars, on public transportation, while working and when recreating.

Recent research has shown that while increasing portions of the population will respond to web requests on their smartphone, the small screen sizes present significant problems. Considerable research summarized elsewhere (Dillman, Hao and Millar 2016) has revealed that the proportion of smartphone responses has increased. In addition, it is difficult to ask many types of questions that seemed to work well in other survey modes. For example, Sarraf, Brooks, Cole and Wang (2015) have shown that the common question format of the item-on-the left with answer categories horizontally displayed to the right and the four-point scale placed below it, resulted in early abandonment of the response process and a dramatic increase in missing responses. In a later set of experiments, Barlas and Thomas (2016) have demonstrated the benefit of shortening scalar questions. These works bring into question the advisability of asking seven point fully

labeled scales, often favored in the past as ideal for interview surveys. Work by Stern, Sterrett and Bilgen (2016) suggest that grids – in which a general question establishing a set of response categories is followed by lists of items requiring an answer to each, a staple of paper and web questionnaires, are not an acceptable visual layout for smartphones.

An excellent review of the available research by Couper et al. (2017) concludes that questionnaires completed on mobile phones have lower response rates, higher breakoff rates, and longer completion times than do web surveys on personal computers. The authors note that part of the reason for these persistent problems may be that surveyors have not yet succeeded in optimizing design for mobile phones. Another factor that contributes to these problems may be competing demands for attention from smartphones and other activities as people are going about the daily rounds of life.

One of the challenges associated with designing for smartphones is maintaining unified question construction across all survey modes. This problem could be particularly acute when respondents in well-established surveys find that previously used question structures, wordings, and visual layouts are unilaterally changed for smartphone use. This challenge has been pointed out by Mistichelli, Eanes and Horwitz of the U.S. Census Bureau (2015). It's not yet clear whether survey designers are willing to change long standing ways of asking questions, (e.g., attitude questions with fewer categories and asking items-in-a series as individual items rather than a list of items introduced with a question that applies to the entire group of items being rated). If unified mode construction is to be used on smartphones, the needs of such devices are likely to be the major determinant of how items are presented across all modes.

The challenge now facing surveyors with regard to smartphones and mobiles also goes much deeper than how to present questions effectively in less space without the need for horizontal and vertical scrolling. In the early days of surveying, in-person interviewers could by their presence to engage the respondent's full attention. With mail, desktop, laptop, and tablets, it might be expected that respondents would often, if not normally, complete surveys at times they were not likely to be interrupted. Smartphones, by their nature are interruption devices, with the possibility of receiving texts, voice phone calls, emails at any moment, often while moving physically through one's daily activities. Answering some surveys may require consultation of records one does not have access to when away from home, or consultation with another household member, which seems harder to achieve if one tries to complete a questionnaire while on the move. The competition for attention that occurs with such devices might lead a surveyor to encourage a respondent not to fill out the survey on a smartphone and instead ask them to do it on their laptop or from home. The problem with that approach is for significant numbers of people smartphones may be their only computer or the only one they attend to on a daily basis. Also, it seems likely that the more one introduces barriers to answering a questionnaire "now", the less likely people are to answer at all.

Working through these issues is one of the largest challenges facing survey methodologists today. But, on a positive note, when multiple modes of contact are used and multiple ways of responding are offered, it seems easier to guide respondents to the most effective way for them to respond as well as for the success of the survey.

6.3 Sponsor Reluctance to undertake mixed-mode surveys and modify single-mode procedures

An additional peril facing web-push surveys is associated with the stress many organizations face in using multiple modes of survey contact and/or response. Each mode of contact and response requires specialized skills, equipment, and software. In order to be effective, it must also be effectively coordinated to deal with many issues at once, as described elsewhere (Dillman et al. 2014, Chapter 11).

Survey sponsors that have specialized in only one form of data collection, or who want to keep data collection activities simple, may be tempted to avoid the use of second or third modes of data collection. This is not likely to happen when high response rates are required (e.g., a national census) or when a substantial economic incentive exists for pushing early respondents to the web (e.g., Biemer et al. 2015). However, the development of do-it-yourself software has encouraged many surveyors to find ways of using only web data collection. Previous research has suggested that significantly biased results toward greater education and income will be produced if data collection stops with only web responses (Rookey, Hanway, and Dillman 2008; Messer and Dillman 2011). Over time this bias may be reduced, but appears not to have happened yet for general populations. Another source of low response rates and potential bias occurs when surveyors obtain only email addresses for a proposed survey, thus eliminating the potential for prior mail contact that allows inclusion of an incentive for encouraging respondents to respond over the Internet.

Making appropriate changes, even when the need is substantial, takes time. In the 1990's the U.S. Census Bureau developed a pre-notice, paper questionnaire, follow-up postcard strategy for data collection (Dillman, Clark and Sinclair 1995), which was used in the 2000 and 2010 Censuses. After the ACS push-to-web strategy was introduced in 2013, it continued to use this approach. The problem it presented is that the follow-up postcard could not provide password information (visible to anyone who picked up the postcard), thus creating the expectation that they would need to return to the web-request letter for that information. Also, the impression of the sequence of sending a pre-notice informing people they would receive a request to respond (by internet), a second letter asking them to go to the web using the provided information, and then a postcard reminder to follow through seemed unnecessarily laborious. Thus, a new procedure of abandoning the pre-notice and using a letter follow-up was introduced. A test of this procedure by the Census Bureau led to its adoption in August 2015 (Clark and Roberts 2016) and a significant increase of 2.5 percentage points in internet responses and a slight reduction in overall costs.

There are many other issues involved in shifting from single-mode thinking to widespread adoption of web-push surveys that involve multiple modes of response. For example, how do researchers overcome the frustration of willing respondents, who are irritated by being told they will have to wait for that request to come in a few weeks? Also, when telephone numbers are available, a phone call could be used as a follow-up reminder with encouragement, rather than simply trying to interview people over the phone. Experimental testing of these alternatives needs to be done.

6.4 Impacts of new discoveries and innovations

Anticipating the future is difficult. When telephone interviewing was rising towards prominence in the 1970's, personal computers were not yet available. And, virtually no one thought, or even imagined, that

only two decades later the telephone that had been tethered to our homes and workplaces would be carried with us nearly everywhere we went using wireless connections. When Internet surveying began in the 1990's few anticipated that not only would the large and clunky desktops that began occupying people's homes would transition to laptops that people would carry with them from place to place. And, that device would later transition to tablets and smartphones with touch screens – both with far more computing power than their original desktops and laptops.

A recent analysis by Friedman (2016) details the monumental changes in the capabilities and power of personal devices that are increasingly taken for granted by large portions of the worldwide population. He traces these capabilities to the exponential growth of each of five different components of today's computers: 1) integrated circuits that do the computing, 2) memory units that store and retrieve information, 3) networking systems that provide communications within and across computers, 4) the software applications that enable different computers to perform various tasks individually and together, and 5) sensors that detect movement, language, light, sound and other features of the environment and turn it into digitized data. He traces the rapid acceleration in these aspects to the development of the iPhone and related innovations occurring since 2007, and their melding into what he describes as the supernova (or cloud).

These developments were only dimly anticipated, even by many of the innovators who created them. Trying to imagine the future is no easier now than it was in the past. For example, voice activation of computer searches is rapidly replacing the individual tapping and swiping of commands on smartphones. Twenty percent of Google searches on Android-powered handsets in the United States are now input by voice (The Economist 2017). In addition, people can also dictate emails and text messages with reasonable success. Will voice-activated answers be the next wave of development for survey designers? It is easy to imagine one being interviewed by his or her smartphone. And, is it possible that simultaneous translations from one language to another, which can now be done with reasonable success, become common on surveys? But, herein lies a fundamental challenge, described by Friedman – the speed with which human beings and societies can adapt to those changes.

Many potential respondents of interest to surveyors still rely on feature phones, while others are racing madly to adopt the most advanced computing and communication device they find practical. And still others are reluctant to use any computer at all. The differences in people's capabilities and preferences require surveyors to be neither too far ahead nor too far behind where most people are.

This raises the issue of whether web-push methods are simply another transitional phase of survey design that may fade out as quickly as it has risen in prominence. The mixed-mode and tailored design focus that now appears to dominate the thinking of survey designers is recognition of the heterogeneity that exists among populations, whose opinions and behaviors surveyors seek to describe.

For a time it appeared that some surveyors thought the value of mixed-mode surveying was in offering people a choice of which mode they would use to respond to a survey request. However, this is only partly true. The real response power of mixed-mode designs for improving response rates stems from making multiple contacts effectively. Each contact gives an opportunity to provide new information about one's survey request and, in some cases, to reach people who cannot be contacted by other survey modes. When survey response requests are offered by different modes there is often an opportunity to improve coverage

(reaching people who can't be reached by another mode) and get people to attend to persuasive arguments for being a respondent. Also, the sequencing of those contacts may help with motivating people to respond (e.g., use of email augmentation of postal letters that makes it easier to respond).

7 Summary and Conclusion

Web-push data collection that begins with a postal mail request to respond over the Internet is one of the major survey design developments of the early 21st century, now offering promise of faster, less-expensive surveys. Many surveyors have been surprised by today's reliance on an initial postal contact. Although mail surveys had often been used to collect survey data, it was expected by many to disappear with the rise of Internet.

The critical development that encouraged reconsideration of mail contact methods greater use was by Link et al. (2008) and Battaglia et al. (2008). This research showed that residential address lists available from the U.S. Postal Service provided the best sample coverage of U.S. residences and could be used to support effective mail surveys of the general public. This work was encouraged by the strong desire to find alternatives to RDD telephone surveys that faced continually declining response and other challenges.

A series of studies, beginning in 2007, looked for ways to use mail contacts to push householders to the web from these address-based lists. This work focused on combining both Internet and paper responses. It was supported by several years of earlier research on measurement differences across modes that showed responses to web and paper questionnaires were quite similar so long as similar question structures, wordings, and visual layouts were used for both data collection methods. Ten experimental comparisons made in these studies received a web-push response rate of 43% of households, with about 60% of the responses coming over the internet and the remainder being obtained by a mail follow-up (Dillman et al. 2014). Major surveys in several countries have researched and adopted the use of web-push methods that rely on not only web and mail, but now include telephone and/or in-person follow-up in their protocols. The goal is to achieve greater response rates and data quality, which a decade ago were thought to be no longer possible in household surveys.

We are now in an era of tailored design in which different survey designs are used for different survey topics, populations, and survey situations. However, it seems likely that web-push data collection methods will see increased use throughout the industrialized world, as survey sponsors seek to benefit from the low cost of internet data collection in order to lower the overall cost of current surveys.

However such methods face challenges that need attention. One is the risk to surveys and respondents from malware, phishing, and server attacks. Another is the increased reliance on smartphones that may require significant changes in how questions are structured and presented to respondents. In addition, the reluctance of organizations and individuals to accept and master the greater complexity associated with shifting from single mode to mixed-mode surveys is a significant challenge.

The history of surveying over the last 75 years has involved significant transitions from the dominance of in-person interviews, to heavy reliance on voice telephone methods, and now to online and mixed mode surveys. It remains to be seen whether web-push methods – now growing in use as a replacement – have a

lasting presence, or will eventually be replaced with web-only data collection, or with other procedures that remain to be innovated or have not yet been conceived.

References

Anderson, M., and Perrin, A. (2016). 13% of Americans don't use the Internet, Who are they? Available at http://www.pewresearch.org/fact-tank/2016/09/07/some-americans-dont-use-the-internet-who-are-they/. Accessed May 2, 2017.

Australian Bureau of Statistics (2016). Making Sense of the Census. Available at http://www.abs.gov.au/websitedbs/censushome.nsf/home/2016. Accessed May 2, 2017.

Barlas, F.M., and Thomas, R.K. (2016). Good questionnaire design: Best practices in the mobile era. *American Association for Public Opinion Research*, January 19th.

Battaglia, M.P., Link, M.W., Frankel, M.R., Osborn, L. and Mokdad, A.H. (2008). An evaluation of respondent selection methods for household mail surveys. *Public Opinion Quarterly*, 72(3), 459-469.

Biemer, P., Murphy, J., Zimmer, S., Berry, C., Deng, G. and Lewis, K. (2016). A test of Web/PAPI protocols and incentives for the residential energy consumption survey. Unpublished paper presented at Annual Conference of the American Association for Public Opinion Research. May 13th.

Blankenship, A.B. (1977). *Professional Telephone Surveys*. New York: McGraw-Hill Book Company.

Blumberg, S.J., and Luke, J.V. (2017). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January-June 2016.

Brick, J.M., and Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *Annals of the American Academy of Political and Social Science*, 645(1), 36-59.

Brick, J.M., Williams, D. and Montaquila, J.M. (2011). Address-based sampling for subpopulation surveys. *Public Opinion Quarterly*, 75(3), 409-428.

Christian, L.M., and Dillman, D.A. (2004). The influence of symbolic and graphical language manipulations on answers to paper self-administered questionnaires. *Public Opinion Quarterly*, 68, 1, 57-80.

Christian, L.M., Dillman, D.A. and Smyth, J.D. (2008). The effects of mode and format on answers to scalar questions in telephone and Web surveys. In *Advances in Telephone Survey Methodology*, (Eds., J.M. Lepkowski, C. Tucker, J.M. Brick, E.D. de Leeuw, L. Japec, P.J. Lavrakas, M.W. Link and R.L. Sangster). New York: Wiley-Interscience, 250-275.

Christian, L.M., Parsons, N.L. and Dillman, D.A. (2009). Designing scalar questions for Web surveys. *Sociological Methods and Research*, 37(3), 393-425.

City of Sapporo (2015). The Japanese government is conducting a Population Census. Available at https://www.city.sapporo.jp/city/english/news/news201508_1e.html. Accessed October 1, 2016.

Clark, S., and Roberts, A. (2016). Evaluation of August 2015 ACS mail contact strategy modification. *2016 American Community Survey Research and Evaluation Report Memorandum Series ACS16-0RER-13.*

Couper, M.P., Antoun, C. and Mavletova, A. (2017). Mobile Web surveys. In *Total Survey Error in Practice*, (Eds., P.P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, N.C. Tucker and B.T. West). New Jersey: Hoboken.

de Leeuw, E.D. (1992). Data quality in mail, telephone and face-to-face surveys. *TT-Publications Amsterdam.*

de Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics,* 21(2), 233-255.

de Leeuw, E., Villar, A., Suzer-Gurtekin, T. and Hox, J. (2017). How to design and implement mixed-mode surveys in cross National Surveys: Overview and guideline. In *Total Survey Error in Practice*, (Eds., P.P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, N.C. Tucker and B.T. West). New Jersey: Hoboken.

Dillman, D.A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons, Inc.

Dillman, D.A. (2000). *Mail and Internet Surveys: The Tailored Design Method, 2nd Edition*. New York: John Wiley & Sons, Inc.

Dillman, D.A. (2002). Navigating the rapids of change: Some observations on survey methodology in the early 21st century. *Public Opinion Quarterly,* 66(3), 473-494.

Dillman, D.A. (2005). Telephone surveys. In *Encyclopedia of Social Measurement,* (Ed., K. Kempf-Leonard), Volume 3. London, UK: Elsevier Press, 757-762.

Dillman, D.A. (2007). *Mail and Internet Surveys: The Tailored Design Method.* 2007 Update with New internet. Visual and Mixed-mode Guide. New Jersey: Hoboken.

Dillman, D.A., and Christian, L.M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods,* 17(1), 30-52.

Dillman, D.A., and Tarnai, J. (1988). Administrative issues in mixed-mode surveys. In *Telephone Survey Methodology*, (Eds., R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II and J. Waksberg), New York: John Wiley & Sons, Inc., 509-528.

Dillman, D.A., and Edwards, M.L. (2016). Designing a mixed-mode survey. In *The SAGE Handbook of Survey Methodology,* (Eds., C. Wolfe, D. Joye, T.W. Smith and Y.-c. Fu), Sage Publications, Thousand Oaks, CA, 255-268.

Dillman, D.A., Clark, J.R. and Sinclair, M.D. (1995). How prenotice letters, stamped return envelopes, and reminder postcards affect mailback response rates for census questionnaires. *Survey Methodology,* 21, 2, 159-165. Paper available at http://www.statcan.gc.ca/pub/12-001-x/1995002/article/14394-eng.pdf.

Dillman, D.A., Gertseva, A. and Mahon-Haft, T. (2005). Achieving usability in establishment surveys through the application of visual design principles. *Journal of Official Statistics,* 21(2), 183-214.

Dillman, D.A., Hao, F. and Millar, M.M. (2016). Improving the effectiveness of online data collection by mixing survey modes. In *The Sage handbook of Online Research Methods, 2nd Edition,* (Eds., N. Fielding, R.M. Lee and G. Blank). Sage Publications, London, 220-237.

Dillman, D.A., Smyth, J.D. and Christian, L.M. (2014). *Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method, 4th Edition*. New Jersey: Hoboken.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. and Messer, B.L. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research*, 38(1), 1-18.

Dutwin, D., and Lavrakas, P. (2016). Trends in telephone outcomes, 2008-2015. *Survey Practice*, 9(3). Available at http://www.surveypractice.org/.

Edwards, M.L., Dillman, D.A. and Smyth, J.D. (2014). An experimental test of the effects of survey sponsorship on Internet and mail survey response. *Public Opinion Quarterly*, 78(3), 734-750.

Finamore, J., and Dillman, D.A. (2013). How mode sequence affects responses by internet, mail and telephone in the national survey of college graduates. Presentation to European Survey Research Association, Ljubljana, Slovenia, July 18.

Friedman, T.L. (2016). *Thank you for Being Late: An Optimist's Guide to Thriving in the Age of Accelerations*. New York, Farrar: Straus and Giroux.

Groves, R.M., and Kahn, R.L. (1979). *Surveys by Telephone*. New York: John Wiley & Sons, Inc.

Groves, R.M., and Peytcheta, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167-189.

Harter, R., Battaglia, M.P., Buskirk, T.D., Dillman, D.A., English, N., Mansour, F., Frankel, M.R., Kennel, T., McMichael, J.P., McPhee, C.B., Montaquila, J., Yancey, T. and Zukerberg, A.L. (2016). Address-base sampling. *American Association for Public Opinion Research* Task Force Report. Available at http://www.aapor.org/getattachment/Education-Resources/Reports/AAPOR_Report_1_7_16_CLEAN-COPY-FINAL-(2).pdf.aspx, 140 pages.

Hoffman, D.D. (2004). *Visual Intelligence: How we Create What we See*. New York: Norton.

Keeter, S., Miller, C., Kohut, A., Groves, R.M. and Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly*, 64(2), 125-148.

Kerlinger, F.N. (1965). *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston.

Krosnick, J.A., and Alwin, D.F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51(2), 201-219.

Link, M.W., Battaglia, M.P., Frankel, M.R., Osborn, L. and Mokdad, A.H. (2008). A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for General Population Surveys. *Public Opinion Quarterly*, 72(1), 6-27.

Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I. and Vehovar. V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, 50(1), 79-104.

McMaster, H.S., LeardMann, C.A., Speigle, S. and Dillman, D.A. (2016). An experimental comparison of web-push vs. paper-only survey procedures for conducting an in-depth health survey of military spouses. BMC Medical Research Methodology.

Messer, B.L. (2012). Pushing households to the web: Results from Web+mail experiments using address based samples of the general public and mail contact procedures. Ph.D. Dissertation. Washington State University, Pullman.

Messer, B.L., and Dillman, D.A. (2011). Surveying the general public over the Internet using address-based sampling and mail contact procedures. *Public Opinion Quarterly*, 75(3), 429-457.

Messer, B.L., Edwards, M.L. and Dillman, D.A. (2012). Determinants of item nonresponse to Web and mail respondents in three address-based mixed-mode surveys of the general public. *Survey Practice*, 5(2), 1-9. Paper available at http://wwww.surveypractice.org/.

Millar, M.M., and Dillman, D.A. (2011). Improving response to Web and mixed-mode surveys. *Public Opinion Quarterly*, 75(2), 249-269.

Mistichelli, J., Eanes, G. and Horwitz, R. (2015). Centurion: Internet Data Collection and Responsive Design. Presentation to Federal Economic Statistics Advisory Committee, June 12.

Mohorko, A., de Leeuw, E. and Hox, J. (2013). Coverage bias in European telephone surveys: Developments of landline and mobile phone coverage across countries and over time. *Survey Methods: Insights from the Field*. Retrieved from http://surveyinsights.org/?p=828.

Nathan, G. (2001). Telesurvey methodologies for household surveys – A review and some thoughts for the future? *Survey Methodology*, 27, 1, 7-31. Paper available at http://www.statcan.gc.ca/pub/12-001-x/2001001/article/5851-eng.pdf.

O'Muircheartaigh, C., English, N. and Eckman, S. (2007). Predicting the relative quality of alternative sampling frames. *2007 Proceedings of the Survey Research Methods Section*, American Statistical Association, [CD ROM], Alexandria, VA: American Statistical Association.

Palmer, S.E. (1999). *Vision Science: Photons to Phenomenology*. London: Bradford Books.

Parten, M. (1950). *Surveys, Polls and Samples*. New York: Harper and Brothers.

Pew Research Center (2012). Assessing the representativeness of public opinion surveys. Available at http://www.people-press.org/2012/05/15/assessing-the-representativeness-of-public-opinion-surveys/. Accessed October 24, 2016.

Redline, C.D., and Dillman, D.A. (2002). The influence of alternative visual designs on respondents' performance with branching instructions in self-administered questionnaires. In *Survey Nonresponse*, (Eds., R. Groves, D. Dillman, J. Eltinge and R. Little), New York: John Wiley & Sons, Inc.

Redline, C.D., Dillman, D.A., Dajani, A. and Scaggs, M.A. (2003). Improving navigational performance in U.S. census 2000 by altering the visual languages of branching instructions. *Journal of Official Statistics*, 19(4), 403-420.

Roberts, C., Joye, D. and Staehli, M.E. (2016). Mixing modes of data collection in Swiss social survey: Methodological report of the LIVES-FORS mixed mode experiment. Lives working paper 2016.48. Swiss National Centre of Competence in Research, a research instrument of the Swiss National Science Foundation.

Rookey, B.D., Hanway, S. and Dillman, D.A. (2008). Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? *Public Opinion Quarterly*, 72(5), 962-984.

Sarraf, S., Brooks, J., Cole, J. and Wang, X. (2015). What is the impact of smartphone optimization on long surveys? Presentation to American Association for Public Opinion Research Annual Conference, Hollywood, FL, May 16.

Smyth, J., Christian, L.M. and Dillman, D.A. (2008). Does 'Yes or No' on the telephone mean the same as check-all-that-apply on the Web? *Public Opinion Quarterly*, 72(1), 103-111.

Smyth, J.D., Dillman, D.A., Christian, L.M. and McBride, M. (2009). Open-ended questions in Web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly*, 73(2), 325-337.

Smyth, J.D., Dillman, D.A., Christian, L.M. and O'Neill, A.C. (2010). Using the Internet to survey small towns and communities: Limitations and possibilities in the early 21st century. *American Behavioral Scientist*, 53(9), 1423-1448.

Smyth, J.D., Dillman, D.A., Christian, L.M. and Stern, M.J. (2006). Comparing check-all and forced-choice question formats in Web surveys. *Public Opinion Quarterly*, 70(1), 66-77.

Statistics Canada (2016). 2016 Census of Population collection response rates. Available at http://www12.statcan.gc.ca/census-recensement/2016/ref/response-rates-eng.cfm. Accessed October 24, 2016.

Statistics Japan (2015). Almost 20 million households responded online in the 2015 Population Census of Japan. Available at http://www.stat.go.jp/english/info/news/20151019.htm. Accessed October 1, 2016.

Stern, M., Sterrett, D. and Bilgen, I. (2016). The effects of grids on Web surveys completed with mobile devices. *Social Currents*, 3(3), 217-233.

Sullivan, M., Leong, C., Churchwell, C. and Dillman, D.A. (2015). Measurement and Cost Effects of Pushing Household Survey Respondents to the Web for Surveys of Electricity and Gas Customers in the United States. Unpublished paper presented to the European Survey Research Association, Reykjavik, Iceland, July 16th.

Tancreto, J. (2012). 2011 American Community Survey Internet Tests: Results from First Test in April 2011. #ACS12-RER-13-R2. 2012 American Community Survey Research and Evaluation Report Memorandum Series, June 25th.

Tarnai, J., and Dillman, D.A. (1992). Questionnaire context as a source of response differences in mail and telephone surveys. In *Context Effects in Social and Psychological Research*, (Eds., N. Schwarz and S. Sudman), New York: Springer Verlag, Inc. 115-129.

The Economist (2017). Now we're talking: Voice technology is making computers less daunting and more accessible. January 7th – 13th, 422 (No. 9022), 9.

Thomas, R., and Barlas, F. (2016). It's a Small Screen After All: Improving Measurement in an Ever-changing Online Survey World. GFK Webinar, September 27th.

Tourangeau, R. (2017). Mixing modes: Tradeoffs among coverage, nonresponse and measurement error. In *Total Survey Error in Practice*, (Eds., P.P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, N. Clyde Tucker and B.T. West), New Jersey, Hoboken: John Wiley & Sons, Inc.

Tourangeau, R., Couper, M.P. and Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly,* 68(3), 368-393.

United Kingdom Cabinet Office (2016). Consultation Response: Community Life Survey: Development and implementation of online survey methodology for future survey years. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/539111/community_life_survey_consultation_response_final.pdf.

United States Census Bureau (2014). American Community Survey Design and Methodology, Version 2.0. Available at http://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html. Accessed October 15, 2016.

Ware, C. (2004). *Information Visualization: Perception for Design, 2nd Edition*. Karlsruhe, West German: Morgan Kaufman.

Williams, D., Brick, J.M., Montaquila, J.M. and Han, D. (2014). Effects of screening questionnaires on response in a two-phase postal survey. *International Journal of Social Research Methodology*, 19(1), 51-67.